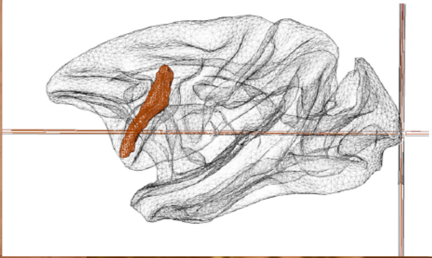
THE EVOLUTION *of* LANGUAGE

Proceedings of the 13th International Conference



Editors
Andrea Ravignani
Chiara Barbieri
Molly Flaherty
Yannick Jadoul
Ella Lattenkamp
Hannah Little
Mauricio Martins
Katie Mudd
Tessa Verhoef

EVOLANG XIII

The Evolution of Language

Proceedings of the 13th International Conference on the Evolution of Language (EvoLang13)

Editors:

Andrea Ravignani
Chiara Barbieri
Molly Flaherty
Yannick Jadoul
Ella Lattenkamp
Hannah Little
Mauricio Martins
Katie Mudd
Tessa Verhoef

EvoLang 13 Scientific Committee
April 14, 2020

Published by:
The Evolution of Language Conferences

Max Planck Institute for Psycholinguistics
Wundtlaan 1
6525 XD Nijmegen
The Netherlands

The Evolution of Language
Proceedings of the 13th International Conference on the Evolution of Language (EvoLang13)

Cover photography and figures by Yannick Becker, Koen de Reus, Sabrina Engesser, Madza Farias-Virgens, Stefan Hartmann, Vanessa Ferdinand, Harun Siljak, Rabia Ergin, Mauricio Martins, and Stephanie White, licensed under a Creative Commons BY-ND 4.0 license.

ISSN 2666-917X
doi:10.17617/2.3190925

Due to the COVID-19 pandemic, this conference could not take place as planned in April 2020. Below you can find the original preface, as it was written before the conference had to be cancelled.

Preface

The 13th International Conference on the Evolution of Language (EvoLang13) was held in Brussels, Belgium from the 14th – 17th of April 2020, organised by Bart de Boer's research group at the Artificial Intelligence Lab of the Vrije Universiteit Brussel. As always, the success of the event is, among other things, facilitated by the team effort and synergy of four bodies: the permanent committee, the local organizers, the scientific committee, and the panel of reviewers.

The *permanent committee* (p.v), headed by Erica Cartimill and Simon Kirby, has always been keen to provide advice and support when needed. Many members of the permanent committee are our mentors or peers, and they keep making the world of language sciences a better place.

The *local organizers*, Bart de Boer, Yannick Jadoul, Katie Mudd, Ross Towns, Marnix Van Soom, and myself have worked hard to make the conference possible. Kudos to this cohesive team for turning an 'EvoLang in Brussels' idea into reality.

The *scientific committee* was in charge of editing and reviewing all contributed abstracts and papers. Once again, the EvoLang XIII scientific committee shows a strong involvement of early career researchers in the evolution of language community. The committee this year spanned several countries and institutions, featuring members at various stages including PhD, postdoc and early PI. By building on the expertise of existing members and recruiting new ones, we really tried to achieve diversity of scientific backgrounds. The fields and perspective covered by the scientific committee span, among other things: developmental psychology, scientific communication, classical linguistics, speech sciences, field research (both in humans and other species), cognitive neuroscience, gesture, artificial intelligence, computational modelling, genetics, anthropology, acoustics, and music cognition.

This iteration of the conference received many high quality submissions, making the hard work of our *reviewers* (p.vi) especially important. Given the quality of submissions we received, the cutoff threshold for acceptance has been particularly high this year. This volume contains many contributions from various disciplines: syntax, semantics, speech sciences, (developmental) psychology, genetics, bioacoustics, anthropology, animal behaviour, and historical linguistics.

The local organising committee, the reviewers, the scientific committee and the permanent committee all voluntarily contribute their time. Thanks to my fellow editors and committee members for bearing with me and acting in coordination to ensure we did the best possible job we could, considering our other commitments.

EvoLang XIII in Brussels marks a number of firsts (we think) in the history of the conference. There is International Sign Language interpreting for one track of the parallel session and at plenary sessions. The pre-conference workshops take one full day, instead of half a day as before; more space needed for, and allocated to, this part of the conference which is organized bottom-up can be seen as a sign of maturity for the field. This edition, three members of the scientific committee are also among the local organizers, which has proved particularly useful for coordination and communication purposes.

On a personal note, this edition marks my 10-years anniversary with the Language Evolution community. Ten years ago, I was a master student in a completely unrelated discipline attending EvoLang 8 in 2010 'just for fun' (to each their own); at this EvoLang in Utrecht I got to discover and fell in love with a fantastic field of study. Since then, it has been an honour and privilege to be welcomed and hopefully contribute to this wonderful community. I sincerely hope that *you*, a student passionate about language evolution who is reading this, will have fun organizing EvoLang in 2030!

- Andrea Ravignani

Central Organizing Committee

- Andreas Baumann
- Rudolf Botha
- Christine Cuskley
- Erica Cartmill
- Jean-Louis Dessalles
- Ramon Ferrer i Cancho
- Tecumseh Fitch
- Jim Hurford
- Simon Kirby
- Chris Knight
- Heidi Lyn
- Luke McCrohon
- Kazuo Okanoya
- Thom Scott-Phillips
- Andrea Ravignani
- Nikolaus Ritt
- Kenny Smith
- Maggie Tallerman
- Natalie Uomini
- Sławomir Wacewicz
- Przemysław Żywiczyński

Panel of Reviewers

Natasha Abner
Ine Alvarez van Tussenbroek
Stephen Anderson
Natalia Aralova
John Archibald
Inbal Arnon
Rie Asano
Mark Atkinson
Chiara Barbieri
Lluís Barceló-Coblijn
Andreas Baumann
Ted Bayne
Teresa Befer
Antonio Benitez-Burraco
Aleksandrs Berdicevskis
Melissa Berthet
Katrien Beuls
Richard Blythe
Cedric Boeckx
Francesca Bonalumi
Monika Boruta-Zywiczynska
James Brand
Erica A. Cartmill
Alexandre Celma-Miralles
Michael Collins
Fred Coolidge
Marie Coppola
Jennifer Culbertson
Chris Cummins
Christine Cuskley
Emily Davis
Bart de Boer
Marianne de Heer Kloots
Koen de Reus
Connie de Vos
Dan Dediu
Peter Dekker
Didier Demolin
Jean-Louis Dessalles
Mark Dingemanse
Natalia Drozhashchikh
Michael Dunn
Mark Ellison
Karen Emmorey
Sabrina Engesser
Caleb Everett
Stephanie Evers
Madza Farias-Virgens
Livio Favaro
Nicolas Fay
Vanessa Ferdinand
Ahana Fernandez
Ramon Ferrer-i-Cancho
Piera Filippi
Adam Fishbein
Molly Flaherty
Stella Frank
Marlen Fröhlich
Koji Fujita
Richard Futrell
Marco Gamba
Maxime Garcia
Timothy Genter
Chris Golston
Nathalie Gontier
Tomás Goucha
Simon Greenhill
Jack Grieve
Callum Hackett
Harald Hammarstrom
Stefan Hartmann
Robert Hawkins
Laurence Henry
Dieter Hillert
Nienke Hoeksema
Aritz Irurtzun
Jessica Ivani
Yannick Jadoul
Niklas Johansson
Sverker Johansson
Tamar Johnson
Anna Jon-And
Andres Karjus
Dimitar Kazakov
Vera Kempe
Rachel Kindellan
Adam King
Simon Kirby
Robert Kluender
Chris Knight
Mirjam Knörnschild
Ski Krieger
Martin Kuhlwilm
Travis LaCroix
Adriano Lameira
Ella Lattenkamp
Li Lei
Ryan Lepic
Shiri Lev-Ari
Roger Levy
Simon Levy
Mattis List
Hannah Little
Gary Lupyan
Heidi Lyn
Vini Macuch Silva
Anna Mai
Michiru Makuuchi
Dan Mann
Konstantina Margiotoudi
Alexander Martin
Mauricio Martins
Pedro Martins
Daiki Matsumoto
Theresa Matzinger
Adrien Meguerditchian
Janine Mengede
Jérôme Michaud
Ashley Micklos
Gregory Mills
Helena Miton
Scott Moisik
Marie Montant
Kathryn Montemurro
Steven Moran
Juan Carlos Moreno Cabrera
Hope Morgan
Juan Moriano
Yasamin Motamedi
Alba Motes Rodrigo
Katie Mudd
Roland Muhlenbernd
Elliot Murphy
Thomas Müller
Savithry Namboodiripad
Jens Nevens
Dillon Niederhut
Iris Nowenstein
Jonas Nölle
Thomas O'Rourke
Kazuo Okanoya
Mikhail Ordin
Aliki Papa
Richard Parncutt
Louise Peckre
Marcus Perlman
Juan Perote-Peña
Michael Pleyer
Piotr Podlipniak
Leona Polyanskaya
Ljiljana Progovac
Andrea Ravignani
Limor Raviv
Nikolaus Ritt
Gareth Roberts
Sean Roberts
Hettie Roebuck
Joana Rossello
Tim Sainburg
Carmen Saldana
Daniela Sammler
Marcelo Sanchez
Sebastian Sauppe
Oliver Sayeed
Sabrina Schalz
Constance Scharff
Adam Schembri
Robert Scholz
Marieke Schouwstra
Thom Scott-Phillips
Yoshimasa Seki
Arthur Semenuks
Dor Shilton
Catriona Silvey
Katie Slocombe
Kenny Smith
Kees Sommer
Göran Sonesson
Matt Spike
Stefanie Sturm
Justin Sulik
Masanori Takezawa
Shira Tal
Monica Tamariz
Carel ten Cate
Claudio Tennie
Bill Thompson
Charles Torres
Valeria Torti
Ross Towns
Genta Toya
Rory Turnbull
Anastasia Ulicheva
Freek Van de Velde
Paul Van Eecke
Carel van Schaik
Olga Vasileva
Tessa Verhoef
Sonja C. Vernes
Jan Verpooten
Slawomir Wacewicz
Joel Wallenberg
Stuart Watson
Andrew Wedel
James Winters
Marieke Woensdregt
Mei-Shin Wu
Ya-Ning Wu
Michael Yartsev
Hiroto Yonenoh
Eva Zehentner
May Zhang
Qing Zhang
Przemek Zywiczynski

Contents

Preface iii

Plenary Talks **1**

Why are sign languages iconic? 2
Carol Padden

Improvisation: A Fundamental Process in Language Emergence 3
Marieke Schouwstra

Oral Traditions and Language Evolution: The Plot Thickens . . 5
Jamie Tehrani

Understanding Bat Vocal Learning to Gain Insight Into Speech and Language 6
Sonja C. Vernes

Refereed Contributions **7**

Evolution of Language Syntax and Musical Rhythm: Flexible Motor and Cognitive Control 8
Rie Asano

Early Left Brain Asymmetry in New-Born Baboons: What Does It Tell Us (Or Not) About the Evolution of the Language-Ready Brain? 12
Yannick Becker, Julien Sein, Lionel Velly, Laura Giacomino, Luc Renaud, Marie Dumasy, Pascaline Boitelle, Romain Lacoste, Jean-Luc Anton, Bruno Nazarian, Adrien Meguerditchian

Older English Words Are More Polysemous 14
Aleksandrs Berdicevskis

Primate Proto-Vowels and the Early Emergence of Speech 22
Louis-Jean Boë, Thomas R. Sawallis

The Advantage of Extreme Meanings in Cultural Evolution ... 24
Fausto Carcassi, Marieke Schouwstra, Simon Kirby

De-Coding the Evolution of Communication: How the Code Model Systematically Biases Studies of Animal Communication 27
Erica A. Cartmill

Did Post-Neolithic Changes in Bite Configuration Impact Speech? A New Approach to the Question 30
Sihan Chen, Caleb Everett

Testing the Language-Throwing Hypothesis 38
Michael Collins

Phonological Cues to Semantic Class Membership Across Hundreds of Languages 41
Pablo Contreras Kallens, Morten H. Christiansen

Evolutionary Perspectives on Speech and Information Rates ... 44
Christophe Coupé, François Pellegrino, Yoon Mi Oh, Dan Dediu

The Co-Evolution of Communication and Cooperation: Lewis Signalling Games in Social Dilemma Scenarios 47
Megan Cross, Amy Lumsden, Matthew Spike

A Game-Based Framework for Studying the Emergence and Dynamics of Shared Conventions 50
Christine Cuskley, Simon Kirby

The Generality and Origins of Form-To-Meaning Iconicity..... 53
Charles Davis, Hannah Morrow, Margherita De Luca, Gary Lupyan

Analysis of Mutation and Fixation for Language 56
Bart de Boer, Bill Thompson, Andrea Ravignani, Cedric Boeckx

Rhythmic Perception, Production and Interactivity in Harbour and Grey Seals.. 59
Marianne de Heer Kloots, Daryll Carlson, Maxime Garcia, Sonja Kotz, Alice Lowry, Laetitia Poli-Nardi, Koen de Reus, Ana Rubio-García, Marlene Sroka, Mila Varola, Andrea Ravignani

Relationships Between Vocal Ontogeny and Vocal Tract Anatomy in Harbour Seals (*Phoca vitulina*) 63
Koen de Reus, Daryll Carlson, Yannick Jadoul, Alice Lowry, Stephanie Gross, Maxime Garcia, Anna Salazar Casals, Ana Rubio-García, Caroline Haas, Bart de Boer, Andrea Ravignani

Vocal Tract Anatomy as a Time Machine: What Can We Infer About the “Deep” Linguistic Past? 67
Dan Dediu, Scott R. Moisik

A Comparative Approach to the Laryngeal Sound Source of Primates: Implications for the Evolution of Speech 71
Didier Demolin

Is There an Iconic Motivation for the Morphophonological Distinction Between Noun-Verb Pairs in ASL? 74
Karen Emmorey, Jennie Pyers

Open Compositionality in Pied Babbler Call Combinations 81
Sabrina Engesser, Amanda R. Ridley, Stuart K. Watson, Sotaro Kita, Simon W. Townsend

Community Structure Affects Convergence on Uniform Word Orders: Evidence From Emerging Sign Languages 84
Rabia Ergin, Limor Raviv, Ann Senghas, Carol Padden, Wendy Sandler

Linguistic Complexity Increases as a Function of Attractiveness in Intersexual Communication. Tentative Experimental Support 87
Charlotte Essers, Freek Van de Velde

Color Salience, Color Term Evolution, and the Issue of Relatedness 90
Caleb Everett, Colleen O’Brien

Evolutionary Mechanisms in the Bengalese Finch’s Song: Parallels and Implications for the Study of Human Speech 93
Madza Farias-Virgens, Terrence Deacon, Kazuo Okanoya, Stephanie White, Emilia Huerta-Sanchez

Fundamental Flaws in Minimalist Accounts of Language: Removing Barriers to an Understanding of Language Evolution . 95
Andrew Feeney

Phoneme Inventory Size and the Language System 98
Gertraud Fenk-Oczlon

Transmission Chains and Individuals Place Different Constraints on the Evolution of Category Systems 102
Vanessa Ferdinand, Amy Perfors, Charles Kemp

Limitations in Avian Perception of Vocal Sequences: Implications for Language Evolution 105
Adam Fishbein, Nora Prior, Gregory Ball, Robert Dooling

An Emergent Language Becomes Smaller as It Evolves: New Evidence From Motion Tracking in Nicaraguan Sign Language 108
Molly Flaherty, Asha Sato, Simon Kirby

Natural Population Growth Can Cause Language Simplification 111
Stella Frank, Kenny Smith

Multimodal and Multiplex Communication in Wild and Captive Orang-Utans (*Pongo abelii/pygmaeus*) 122
Marlen Fröhlich, Carel P. van Schaik

Cognitive Constraints Built Into Formal Grammars: Implications for Language Evolution 125
Carlos Gomez-Rodriguez, Morten H. Christiansen, Ramon Ferrer-i-Cancho

Revisiting the Uniformitarian Hypothesis: Can We Detect Recent Changes in the Typological Frequencies of Speech Sounds? 128
Eitan Grossman, Elad Eisen, Dmitry Nikolaev, Steven Moran

The Emergence of Compositional Languages for Numeric Concepts Through Iterated Learning in Neural Agents 131
Shangmin Guo, Yi Ren, Serhii Havrylov, Stella Frank, Ivan Titov, Kenny Smith

A New Hope (For Common Ground)? Convergences Between Biolinguistic and Usage-Based Approaches to Language and Its Evolution 141
Stefan Hartmann, Michael Pleyer

Gradualism in the Evolution of Ostensive Communication 144
Christophe Heintz, Thom Scott-Phillips

Could Vocal Communities of Territorial Songbirds Constitute a Pseudo-Social System? Testing Dialects as Potential Community Markers in Social Versus Territorial Congeneric Songbird Species 155
Laurence Henry, Maël Leroux, Léa Lange, Enya Quéré, Isabelle Naas, Adrian Craig, Martine Hausberger

Showing and Seeing: Enactment in British Sign Language Conversations 159
Gabrielle Hodge, Sannah Gulamani, Kearsy Cormier

Neuroanatomy of the Grey Seal Brain: Bringing Pinnipeds Into the Neurobiological Study of Vocal Learning 162
Nienke Hoeksema, Stella Villanueva, Janine Mengede, Anna Salazar Casals, Ana Rubio-García, Branislava Curcic-Blake, Sonja C. Vernes, Andrea Ravignani

Bats and the Comparative Neurobiology of Vocal Learning 165
Nienke Hoeksema, Maximilian Wiesmann, Amanda Kiliaan, Peter Hagoort, Sonja C. Vernes

Evolutionary Forces in the Development of the English Perfect Construction 168
Michio Hosaka, Shimpei Okuda, Kazutoshi Sasahara

The Effects of Interaction on Establishing Common Ground in Communication 171
Elliot Hugget, Paola Peña, Justin Sulik, Matthew Spike

Communicative Context and the Evolution of Language: Change Over Time in the Rhetorical Structure of British Patent Specifications, 1711 – 2011 174
Grieve Jack, Nicholas Groom

A Phylogenetic Study of Sino-Tibetan Kinship Evolution 177
Ting Ji, May Zhang, Thomas Currie, Andrew Meade, Mark Pagel, Ruth Mace

Evolutionary Rates and Directions of Word Order Change Across Main and Subordinate Clauses in Indo-European 181
Yingqi Jing, Paul Widmer, Balthasar Bickel

The Underlying Dimensions of Sound-Color Correspondences Reveal Typologically and Evolutionary Grounded Linguistic Primitives 186
Niklas Johansson, Andrey Anikin

Pointer Evolution Points to Gradual Evolution of Hierarchical Complexity 189
Sverker Johansson

Assessing Measures of Morpholoigcal Complexity as Predictors of Learning by Neural Networks and Humans 197
Tamar Johnson, Kexin Gao, Jennifer Culbertson, Hugh Rabagliati, Kenny Smith

Minimal Prerequisits for Processing Language Structure: A Model Based on Chunking and Sequence Memory 200
Anna Jon-And, Jérôme Michaud

Illustrating the Creative Aspects of Sound Symbolism: Implications for Theories of Language Evolution 209
Keith Jones, Marcus Perlman

Modelling Cultural Evolution of Pragmatic Communication When Language Co-Develops With Perspective-Taking 220
Anna Kapron-King, Simon Kirby, Marieke Woensdregt

Manipulating Learning Constraints and Signal Features in Iterated Language Learning 223
Vera Kempe, Sheila Cunningham, Nikolay Panayotov, Nicolas Gauvrit, Monica Tamariz

The Role of Simplicity in Word Order Harmony 225
Rachel Kindellan, Danielle Naegeli, Kenny Smith, Jennifer Culbertson

The Lexicon as a Huffman Code: Words Are Structured for Probabilistically Balanced Contrasts 234
Adam King, Andrew Wedel

Cumulative Cultural Evolution of Systematic Structure and Grammatical Complexity in Humans and Baboons 237
Simon Kirby, Vanessa Ferdinand, Kenny Smith, Joël Fagot, Nicolas Claidière

The Role of Ritual in the Evolution of Language and the Emergence of Recursion 240
Robert Kluender, Emily Davis

Impact of Auditory Feedback on Bat Vocal Development 249
Ella Lattenkamp, Meike Linnenschmidt, Eva Mardus, Sonja C. Vernes, Lutz Wiegrebe, Michael Schutte

Using Spatial Visualizations and Real-World Social Networks to Understand Language Evolution and Change 252
Li Lei, Limor Raviv, Phillip Alday

Call Combinations in Chimpanzees 255
Maël Leroux, Katie Slocombe, Anne Schel, Klaus Zuberbühler, Simon W. Townsend

Communities of Different Size Create Different Categorization Systems 257
Shiri Lev-Ari

A Point Is Not a Point Is Not a Point: Re-Interpreting Three Basic Kinds of Pointing Comprehension 260
Heidi Lyn, Jennie Christopher

Does a Universal Hierarchical Structure Underlie Word Order Typology? 263
Alexander Martin, Patrick Kanampiu, David Adger, Klaus Abels, Jennifer Culbertson

Specifying the "Vocal" in Vocal Learning 266
Pedro Tiago Martins, Cedric Boeckx

Towards a Comparative Approach to (Syntactic) Hierarchy 269
Daiki Matsumoto

Aesthetic Perception of Linguistic Features as a Potential Factor in Language Change 272
Theresa Matzinger, Eva Specker

The Symbolic Capacity as a Basis of Human Cognition in Evolution and Development 276
Larissa Mendoza Straffon

Studying the Genetic Bases of Vocal Learning in Bats 280
Janine Mengede, Paolo Devanna, Stephen Hörpel, Uwe Firzlaff, Sonja C. Vernes

A Game Theoretic Perspective on the Utterance Selection Model for Language Change 283
Jérôme Michaud

Human Language as a (Unique) Combination of Inherited (And Shared) Domain-General Cognitive Abilities 291
Marie Montant, Raphaelle Malassis, Tiphaine Medam, Joël Fagot

Genetic Networks Under Human-Specific Regulation Reveals Clues About the Evolution of the Modern Language-Ready Brain 294
Juan Moriano Palacios, Alejandro Andirkó, Cedric Boeckx

The Seeds of the Noun-Verb Distinction in the Manual Modality: The Role of Improvisation and Interaction 296
Yasamin Motamedi, Kathryn Montemurro, Natasha Abner, Molly Flaherty, Simon Kirby, Susan Goldin-Meadow

Locally Restricted Behaviors: In Search of Potential Great Ape Culture Dependent Traits 299
Alba Motes Rodrigo, Claudio Tennie

How Does Social Structure Shape Language Variation? A Case Study of the Kata Kolok Lexicon 302
Katie Mudd, Hannah Lutzenberger, Connie de Vos, Paula Fikkert, Onno Crasborn, Bart de Boer

A Cultural-Evolutionary Model of Politeness Strategies and Reputation 305
Roland Mühlenbernd, Slawomir Wacewicz, Przemyslaw Zywiczynski

Structural Similarities Between Natural Languages and Artificial Languages in the Domain of Color and Their Impact on Communication: A Large-Scale Online Study 308
Thomas Müller, James Winters, Tiffany Morisseau, Ira Noveck, Olivier Morin

Learning General, Adaptive and Human-Interpretable Concepts Through Situated Interactions 311
Jens Nevens, Paul Van Eecke, Katrien Beuls

Emergence of Communication Under Conflict of Interest 314
Michael Noukhovitch, Travis Lacroix, Angeliki Lazaridou, Aaron Courville

Iconicity in Sign Grounding: Representation or Disambiguation? 318
Jonas Nölle, Riccardo Fusaroli, Kristian Tylén

Does Environment Shape Spatial Language? A Virtual Reality Experiment 321
Jonas Nölle, Simon Kirby, Jennifer Culbertson, Kenny Smith

Glutamate Receptors Implicated in (Self-)Domestication Regulate Dopaminergic Signaling in Striatal Vocal-Learning Pathways . . . 324
Thomas O'Rourke, Cedric Boeckx

Evolution of Vocal Complexity: A Peak-Shift and Epigenetics Account . . . 328
Kazuo Okanoya

Linguistic Experience and Universal Design of the Auditory System in Speech Rhythm Perception: Implications for Speech Evolution . . . 331
Mikhail Ordin, Leona Polyanskaya, David Maximiliano Gómez, Arthur Samuel

Copying From Observing a Model vs. Listening to Linguistic Instructions: Comparing Children vs. Adults, and Causally Relevant vs. Irrelevant Actions . . . 334
Aliki Papa, Mioara Cristea, Nicola McGuigan, Monica Tamariz

The Obstetric Dilemma and the Origin of Language . . . 337
Richard Parncutt

Pretend Play as a Scaffold for Language Evolution . . . 345
Michael Pleyer

Lexical and Gestural Alignment in Interaction and the Emergence of Novel Shared Symbols . . . 356
Marlou Rasenberg, Mark Dingemanse, Asli Özyürek

Network Structure and the Cultural Evolution of Linguistic Structure: A Group Communication Experiment . . . 359
Limor Raviv, Antje Meyer, Shiri Lev-Ari

The Emergence of Phonology-Like Organization in a Quasi-Phonetic Space . . . 362
Gareth Roberts, Robin Clark

Does Greater Use of Language Promote Greater Conceptual Alignment? . . . 365
Hettie Roebuck, Gary Lupyan

The Evolution of Vocative Morphology as Shaped by the Tune . . . 368
Timo B. Roettger, Márton Sóskuthy

Visual Recursion Develops in Absence of Linguistic Recursion. A Case Report 371
Joana Rosselló, Alexandre Celma-Miralles, Mauricio D. Martins

Long-Range Sequential Dependencies Precede Syntactically-Rich Vocalizations in Humans 374
Tim Sainburg, Anna Mai, Timothy Genter

Rational After All: Changes in Probability Matching Behaviour Across Time in Humans and Monkeys 377
Carmen Saldana, Nicolas Claidière, Joël Fagot, Kenny Smith

Language Discrimination by Large-Billed Crows 380
Sabrina Schalz, Ei-Ichi Izawa

Understanding Verb Directionality in Emerging and Established Sign Languages 390
Adam Schembri, Kearsy Cormier

Simulating Lexical Semantic Change From Sense-Annotated Data 393
Dominik Schlechtweg, Sabine Schulte Im Walde

Learning vs Use in Language Evolution 405
Thom Scott-Phillips, Adam Jowett, Sean Roberts

The Co-Evolution of Value Systems and Institutional Reinforcement: A Micro-Scale Model of Dissemination of Communicative Variants 408
José Segovia-Martín, Monica Tamariz

Phylogenetic Exploration of Language Complexity in Austronesian, Bantu, and Indo-European Language Families 411
Olena Shcherbakova, Simon Greenhill, Hedvig Skirgård

The Communicative Demands of Toolmaking Skill Transmission 414
Dor Shilton

Do Life Histories Shape Vocal Production Learning? A Bird-Based Approach 417
Sara Silvente i Font, Pedro Tiago Martins, Cedric Boeckx

Dominance Style and Vocal Communication in Non-Human Primates 420
Katie Slocombe, Eithne Kavanagh, Sally Street, Felix O. Angwela, Thore J. Bergman, Maryjka B. Blaszczyk, Laura M. Bolt, Margarita Briseño-Jaramillo, Michelle Brown, Chloe Chen-Kraus, Zanna Clay, Camille Coye, Melissa Emery Thompson, Alejandro Estrada, Barbara Fruth, Claudia Fichtel, Marco Gamba, Cristina Giacoma, Kirsty E. Graham, Samantha Green, Cyril Grueter, Shreejata Gupta, Morgan L. Gustison, Lindsey Hagberg, Daniela Hedwig, Katharine M. Jack, Peter M. Kappeler, Gillian King-Bailey, Barbora Kuběnová, Alban Lemasson, David MacGregor Inglis, Zarin Machanda, Andrew MacIntosh, Bonaventura Majolo, Sophie Marshall, Stephanie Mercier, Jérôme Micheletta, Martin Muller, Hugh Notman, Karim Ouattara, Julia Ostnera, Mary S. M. Pavelka, Louise R. Peckre, Megan Petersdorf, Fredy Quintero, Gabriel Ramos-Fernández, Martha M. Robbins, Roberta Salmi, Isaac Schamberg, Oliver Schülke, Stuart Semple, Joan B. Silk, J. Roberto Sosa-Lopéz, Valeria Torti, Daria Valente, Raffaella Ventura, Erica Van De Waal, Anna H. Weyher, Claudia Wilke, Richard Wrangham, Christopher Young, Anna Zanoli, Klaus Zuberbühler, Adriano Lameira

Variation in Mindreading "On the Ground": Comparing Patterns of Mental State Talk in Two Societies 423
Andrew Marcus Smith, H. Clark Barrett, Erica A. Cartmill

Did Partial Hearing Loss Shape Australian Phonology? An Empirical Test Using Spoken Iterated Learning 426
Kenny Smith, Dan Dediu

Cerebellar Forward Models and Their Role in Auditory Processing 429
Stefanie Sturm, Iria SanMiguel

Appeals to 'Theory of Mind' No Longer Explain Much in Language Evolution 433
Justin Sulik, Gary Lupyan

Redundant Morphological Marking Benefits Child Learners 437
Shira Tal, Inbal Arnon

Complexity Can Be Maintained in Small Populations: A Model of Lexical Variability in Emerging Sign Languages 440
Bill Thompson, Limor Raviv, Simon Kirby

Co-Evolution of Culture and Meaning Revealed Through Large-Scale Semantic Alignment 443
Bill Thompson, Sean Roberts, Gary Lupyan

Sexual Dimorphic Chorusing in the Wild Indris 447
Valeria Torti, Chiara De Gregorio, Daria Valente, Anna Zanoli, Longondraza Miaretsoa, Cristina Giacoma, Marco Gamba

Chunking and Associative Learning in Non-Human Primates .. 451
Laure Tosatto, Dezso Nemeth, Joël Fagot, Arnaud Rey

The Evolution of Spatial Devices in Gestural Storytelling 454
Ross Towns, Marcello A. Gómez-Maureira, Kees Sommer, Max van Duijn, Tessa Verhoef

Investigating the Cultural Evolution of the Blombos and Diepkloof Engravings 457
Kristian Tylén, Riccardo Fusaroli, Sergio Gonzalez De La Higuera Rojo, Katrin Heimann, Nicolas Fay, Niels Nørkjær Johannsen, Felix Riede, Marlize Lombard

Gestural Origins of Verb and Noun Encoding in Sign Language Emergence 460
Tessa Verhoef, Ryan Lepic

Did Language Evolve From Indexical Signaling? 463
Till Nikolaus von Heiseler

Language-Specific Constraints on Word Form Predict Segment Information Across the Word 471
Andrew Wedel, Adam King

Modelling the Role of Other-Initiated Repair in Facilitating the Emergence of Compositionality 474
Marieke Woensdregt, Mark Dingemanse

That's (NOT) About the Size of It: Sound-Symbolic Effect of Labels on Size Perception 477
Ya-Ning Wu, Arturs Semenuks

Your Theory of Language Evolution Also Depends on Your View of Evolution 480
Thiago Macek Gonçalves Zahn

Emergence of Pragmatic Reasoning From Least-Effort Optimization 484
Noga Zaslavsky, Jennifer Hu, Roger Levy

Author Index **489**

Plenary Talks

WHY ARE SIGN LANGUAGES ICONIC?

CAROL PADDEN

cpadden@ucsd.edu
University of California, San Diego

The standard answer to this question is that sign languages are iconic because they *can* be, as visual-manual systems. But to ask why they are iconic means to ask why they *need to be.* Knowing that sign languages are iconic does not explain how iconicity works in human language, nor how it works in other manual visual systems that are not sign languages, such as in fingerspelling or in auxiliary sign languages such as Warlpiri Sign Language which is used alongside spoken Warlpiri in the central desert region of Australia. In this talk, I compare sign languages to other visual-manual systems that have been described in the literature. Deaf community sign languages or deaf sign languages constitute one group, though both terms are problematic because they suggest that deaf people are the exclusive developers of or the propagators of sign languages which is not always true. Hearing people play a much larger role in the creation and propagation of sign languages than has been described in the literature. Sign languages are largely *metonymic* systems, referring to some aspect of referents. Manual alphabetic systems that accompany sign languages are iconic, but characteristically of the written characters themselves, not of referents. Finally, Warlpiri Sign Language is called an “auxiliary” or a “secondary” system because it is produced in close relationship with a spoken language. As it is used among hearing speakers of Warlpiri, it appears that its iconicity is of a yet different type. I argue that the differences in iconicity across these different manual systems demonstrate that iconicity is not an independent property of language but is constituted by it. In other words, iconicity and grammar are co-constitutive. I further show, using data from a young sign language used by Bedouins in Israel that, like grammar, iconicity is not immediately available to human language but takes time to emerge in a community of users.

IMPROVISATION: A FUNDAMENTAL PROCESS IN LANGUAGE EMERGENCE

MARIEKE SCHOUWSTRA

Marieke.Schouwstra@ed.ac.uk
Centre for Language Evolution, University of Edinburgh, Edinburgh, UK

All languages have systematic structure, allowing us to communicate an unbounded number of thoughts. Among those who study where this structure comes from, there is increasing recognition that both individual and cultural processes play a role in the evolution of language. But it is not clear how these processes interact. My research tackles this issue by focusing on what happens when languages are created anew. I do this by conducting laboratory experiments in which naive participants improvise to communicate in novel ways, and by comparing my lab results to actual cases of recent language emergence, for example Nicaraguan Sign Language.

One way to elicit improvised utterances is by using silent gesture: participants use their hands and bodies but no speech to convey information. This method has been applied successfully in studies of Basic Word Order, as a way to uncover individual (largely native-language-independent) biases for structuring information about transitive events. In my work I combine silent gesture with experimental techniques from language evolution more broadly, to see what happens to the results of improvisation under the influence of communicative interaction and iterated learning.

A key finding has been that laboratory participants, when they improvise, use structures that are shaped by the semantic properties of the message to be conveyed. The structure of their utterances reflects the way in which they conceptualise the world. I call this tendency naturalness, and show that naturalness, together with iconicity, is an effective bias to get a novel communication system off the ground. However, my experiments show that when multiple utterances are produced and perceived in communicative interaction and when language is transmitted to new generations, this brings new biases to the fore, replacing the biases dominant in improvisation.

In this talk, I will review experimental evidence from different linguistic domains, for instance, Basic Word Order, Noun Phrase structure, and the structure of the lexicon, and give an overview of the biases we observed in improvisation, communicative interaction, and iterated learning.

Observing the birth of linguistic systems in the laboratory is exciting and informative, but it is important to recognise that the idealised setting of the laboratory might miss crucial factors that are important in shaping real emerging languages. Because of this I will discuss ongoing work in which we compare laboratory results concerning basic word order to sentence structures from a young sign language.

Acknowledgements

Marieke Schouwstra is funded by the ESRC, ES/R011869/1.

ORAL TRADITIONS AND LANGUAGE EVOLUTION: THE PLOT THICKENS

JAMIE TEHRANI

jamie.tehrani@durham.ac.uk
Durham Cultural Evolution Research Centre (DCEC)
Department of Anthropology
Durham University

In this talk I will explore the numerous points of contact that exist between studies of language evolution and the transmission of oral traditions, such as folktales, myths and legends. These connections have deep and tangled historical roots, from the work of nineteenth century writers like Augustus Schleicher and the Brothers Grimm, who attempted to reconstruct Proto-Indo-European language and folklore, to Frederic Bartlett's pioneering studies of folktale transmission that helped establish the iterated learning paradigm so widely used in modern-day experimental studies of language evolution. While the two fields grew apart in the latter half of the twentieth century, the emergence of Cultural Evolution has opened up new possibilities for cross-fertilisation. I will discuss the possibilities and prospects for developing these synergies by presenting a series of case studies on issues of mutual interest, including the role of transmission biases/ factors of attraction, cultural diffusion and population histories, and the relationship between demography and cultural complexity.

UNDERSTANDING BAT VOCAL LEARNING TO GAIN INSIGHT INTO SPEECH AND LANGUAGE

SONJA C. VERNES

sonja.vernes@mpi.nl
Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, PO Box 310, Nijmegen, 6500 AH, The Netherlands
Donders Institute for Brain, Cognition and Behaviour, Kapittelweg 29, Nijmegen, 6525 EN, The Netherlands

The comparative approach can provide insight into the evolution of speech, language, and social communication by studying relevant traits in animal systems. Bats are emerging as a model system with great potential to shed light on these processes given their learned vocalisations, close social interactions, and mammalian physiology. Bats are highly social animals that have developed sophisticated vocal and auditory systems for navigation and communication. Furthermore, their small size, amenability to manipulation, and the rich history of studying the neuroethological of echolocation in bats, makes them an excellent system to model the biology underlying vocal learning behaviour. I will discuss the evidence for vocal learning in bats. This encompasses our understanding of the abilities bats have displayed for vocal learning, what is known about the timing and social structure needed for such learning, and current knowledge about the prevalence of the trait across the order. I will highlight approaches being undertaken to understand the neurobiological and genetic mechanisms underlying this complex behaviour, and raise some key questions that should be answered to advance our understanding of the biological encoding and evolution of speech and spoken communication.

Refereed Contributions

EVOLUTION OF LANGUAGE SYNTAX AND MUSICAL RHYTHM: FLEXIBLE MOTOR AND COGNITIVE CONTROL

RIE ASANO*[1,2]

*Corresponding Author: rie.asano@uni-koeln.de
[1]Department of Systematic Musicology, University of Cologne, Cologne, Germany

From theoretical as well as empirical perspectives, there is an increasing number of evidences for a close relationship between language syntax and musical rhythm (e.g., Fitch, 2013; Gordon, Jacobs, Schuele, & McAuley, 2015), although this link seems not to be intuitive and tends to be neglected so far. In the current paper, I indeed argue for the strong link between language syntax and musical rhythm by showing that the same neurocognitive mechanisms implemented in the cortico-basal ganglia-thalamocortical (CBGT) circuits are involved in and necessary for flexible motor and cognitive control crucial for both domains. Moreover, I propose that phylogenetic changes in the CBGT circuits led to the gradual transition from goal-directed, reward-based motor control to more flexible motor and cognitive control.

First of all, several neuroimaging and neuropsychological studies point out critical role of the CBGT circuits for processing musical rhythm and language syntax. Concerning musical rhythm processing, the motor CBGT circuit including the supplementary motor area (SMA) and the putamen is indicated in continually predicting the next beat (Grahn & Rowe, 2013) as well as hierarchical structuring of rhythm (Asano, 2019). In addition, Parkinson's disease (PD) patients show impaired beat-based encoding of rhythm mainly in performing perceptual tasks (for a review, see Leow & Grahn, 2014). The executive CBGT circuit including the prefrontal cortex and the caudate is activated in finding the beat (Kung, Chen, Zatorre, & Penhune, 2013) as well as processing breaches of expectation (Schiffer & Schubotz, 2011).

As for language syntax, morphosyntactic and phrasal syntactic processing such as word order violation processing (Moro et al., 2001) and syntactic ambiguity resolution (Stowe, Paans, Wijers, & Zwarts, 2004) recruit the executive circuit. The putamen (but not the SMA) was indicated in processing phrase structure violation (Friederici, Rüschemeyer, Hahne, & Fiebach, 2003) and increasingly larger constituent structure (Moreno, Limousin, Dehaene, & Pallier, 2018). Moreover, PD, Huntington's disease (HD), and focal basal ganglia patients display problem in inhibiting an overlearned syntactic representation and selecting an alternative (for reviews, see Friederici & Kotz, 2003; Kotz, Schwartze, & Schmidt-Kassow, 2009). Especially, HD patients and even their asymptomatic relatives showed abnormal processing sentences with center-embedding (García et al., 2017). Processing sequences with center-embedding also activates the caudate (Bahlmann, Schubotz, & Friederici, 2008).

That is, the CBGT circuits are involved in and necessary for processing musical rhythm and language syntax although the former primarily relies on the motor circuit and the latter on the executive circuit.

The basal ganglia are known to contribute as a 'control center' working together with the cortico-thalamic pattern generators in both motor and cognitive domains: 1) assisting execution of cortically driven predictable and automatic motor and cognitive patterns; and 2) adapting to unusual circumstances by interrupting and altering the automatically running motor and cognitive representations (Graybiel, 1997; Marsden & Obeso, 1994). Those dual functions of the basal ganglia facilitate flexibility and adaptation in motor and cognitive control. Indeed, motor and cognitive flexibility is crucial for processing musical rhythm and language syntax. For example, in continually predicting the next beat, execution of the current time interval represented in the SMA should be assisted by suppressing the alternative representations. In processing breaches of expectation caused by structural violations, structural ambiguity, and center-embedding, automatic cortical representations should be suppressed to choose an alternative one. Thus, I suggest that flexible control implemented in the CBGT circuits is the common basis of musical rhythm and language syntax.

Goal-directed, reward-based adaptation in motor control through the basal ganglia is wide-spread in animals. However, mice with humanized Foxp2 shows change in the basal ganglia (increase in total dendrite length of the striatal medium spiny neurons) and significantly more rapid switching of their behavioral strategy (Enard, 2011; Scharff & Petri, 2011; Schreiweis et al., 2014). Further, between-species difference in dopaminergic innervation of the caudate is reported (Raghanti et al., 2016). Thus, I claim that phylogenetic changes in the CBGT circuits led to increasingly more flexible motor and cognitive control in humans, making the brain partly ready for processing musical rhythm and language syntax.

References

Asano, R. (2019). *Principled Explanations in Comparative Biomusicology – Toward a Comparative Cognitive Biology of the Human Capacities for Music and Language*. University of Cologne.

Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage, 42*(2), 525–534. https://doi.org/10.1016/j.neuroimage.2008.04.249

Enard, W. (2011). FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. *Current Opinion in Neurobiology*, *21*(3), 415–424. https://doi.org/10.1016/j.conb.2011.04.008

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, *7,* 68. https://doi.org/10.3389/fnsys.2013.00068

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: functional imaging and lesion studies. *NeuroImage*, *20*(SUPPL. 1), S8–S17. https://doi.org/10.1016/j.neuroimage.2003.09.003

Friederici, A. D., Rüschemeyer, S.-A., Hahne, A., & Fiebach, C. J. (2003). The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. *Cerebral Cortex*, *13*(2), 170–177. https://doi.org/10.1093/cercor/13.2.170

García, A. M., Bocanegra, Y., Herrera, E., Pino, M., Muñoz, E., Sedeño, L., & Ibáñez, A. (2017). Action-semantic and syntactic deficits in subjects at risk for Huntington's disease. *Journal of Neuropsychology*. https://doi.org/10.1111/jnp.12120

Gordon, R. L., Jacobs, M. S., Schuele, C. M., & McAuley, J. D. (2015). Perspectives on the rhythm-grammar link and its implications for typical and atypical language development. *Annals of the New York Academy of Sciences*, *1337*(1), 16–25. https://doi.org/10.1111/nyas.12683

Grahn, J. A., & Rowe, J. B. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cerebral Cortex*, *23*(4), 913–921. https://doi.org/10.1093/cercor/bhs083

Graybiel, A. M. (1997). The Basal Ganglia and Cognitive Pattern Generators. *Schizophrenia Bulletin*, *23*(3), 459–469. https://doi.org/10.1093/schbul/23.3.459

Kotz, S. A., Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*, *45*(8), 982–990. https://doi.org/10.1016/j.cortex.2009.02.010

Kung, S.-J., Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2013). Interacting Cortical and Basal Ganglia Networks Underlying Finding and Tapping to the Musical Beat. *Journal of Cognitive Neuroscience*, *25*(3), 401–420. https://doi.org/10.1162/jocn_a_00325

Leow, L.-A., & Grahn, J. A. (2014). Neural Mechanisms of Rhythm Perception:

Present Findings and Future Directions. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing* (pp. 325–338). New York: Springer New York. https://doi.org/10.1007/978-1-4939-1782-2_17

Marsden, C. D., & Obeso, J. A. (1994). The functions of the basal ganglia and the paradox of stereotactic surgery in Parkinson's disease. *Brain, 117,* 877–897.

Moreno, A., Limousin, F., Dehaene, S., & Pallier, C. (2018). Brain correlates of constituent structure in sign language comprehension. *NeuroImage, 167,* 151–161. https://doi.org/10.1016/j.neuroimage.2017.11.040

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., & Fazio, F. (2001). Syntax and the Brain: Disentangling Grammar by Selective Anomalies. *NeuroImage*, *13*(1), 110–118. https://doi.org/10.1006/nimg.2000.0668

Raghanti, M. A., Edler, M. K., Stephenson, A. R., Wilson, L. J., Hopkins, W. D., Ely, J. J., … Sherwood, C. C. (2016). Human-specific increase of dopaminergic innervation in a striatal region associated with speech and language: A comparative analysis of the primate basal ganglia. *Journal of Comparative Neurology, 524*(10), 2117–2129. https://doi.org/10.1002/cne.23937

Scharff, C., & Petri, J. (2011). Evo-devo, deep homology and FoxP2: implications for the evolution of speech and language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1574), 2124–2140. https://doi.org/10.1098/rstb.2011.0001

Schiffer, A.-M., & Schubotz, R. I. (2011). Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Frontiers in Human Neuroscience, 5,* 38. https://doi.org/10.3389/fnhum.2011.00038

Schreiweis, C., Bornschein, U., Burguière, E., Kerimoglu, C., Schreiter, S., Dannemann, M., … Graybiel, A. M. (2014). Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. *Proceedings of the National Academy of Sciences*, *111*(39), 14253–14258. https://doi.org/10.1073/pnas.1414542111

Stowe, L. A., Paans, A. M. ., Wijers, A. A., & Zwarts, F. (2004). Activations of "motor" and other non-language structures during sentence comprehension. *Brain and Language*, *89*(2), 290–299. https://doi.org/10.1016/S0093-934X(03)00359-6

EARLY LEFT BRAIN ASYMMETRY IN NEW-BORN BABOONS: WHAT DOES IT TELL US (OR NOT) ABOUT THE EVOLUTION OF THE LANGUAGE-READY BRAIN?

YANNICK BECKER[1], JULIEN SEIN[2], LIONEL VELLY[2], LAURA GIACOMINO[2], LUC RENAUD[2], MARIE DUMASSY[3], PASCALINE BOITELLE[3], ROMAIN LACOSTE[3], JEAN-LUC ANTON[2], BRUNO NAZARIAN[2]
and ADRIEN MEGUERDITCHIAN[*1,3]

[*]Corresponding Author: adrien.meguerditchian@univ-amu.fr
[1]Laboratoire de Psychologie Cognitive, CNRS, Aix-Marseille Univ, Marseille, France
[2]Institut des Neuroscience de la Timone, CNRS, Aix-Marseille Univ, Marseille, France
[3]Station de Primatologie CNRS, Rousset, France

Language and its typical asymmetric brain organization are considered as unique to *Homo sapiens* evolution (Crow, 2002), suggesting a specific “language-ready” brain dating back to 350 000 ago. Therefore brain lateralization was hypothesized as central to characterize the language-ready brain, a majority of humans showing greater cortical activations in the left hemisphere for most language functions (Vigneau et al. 2006). The specific “language-ready” brain also suggests infant brain is pre-wired to learn languages. At birth preverbal new-born infants are already sensitive to their native vocal prosody and are able to distinguish every phonemes (e.g. Mehler, 1988), suggesting the innate inherited readiness for language acquisition. Similarly than adults, infants also show structural and functional leftward asymmetry of a critical language area, the planum temporale (PT), highlighting the PT asymmetry as a marker for functional language readiness in children development prior to language exposure (e.g. Dubois, 2009). Surprisingly, PT structural asymmetry has been now reported in nonhuman species such as chimpanzees and baboons (e.g. Marie et al., 2018).

The aim of the present study is to investigate the neuroanatomical asymmetries of the PT in new-born nonhuman primates, the baboons (*Papio anubis*). T1 & T2 weighted anatomical images were acquired *in vivo* on a 3T MRI scanner in 32 unwean infant baboons of different ages (from 4 days to 2 months old at the Centre IRM (Institut de Neurosciences de la Timone) before the full maturation of myelin, synapses and cell bodies. Among those baboons, 18 have been scanned a second time when reaching 7 to 10 months of age (i.e., after myelin maturation). Both infant and mother were anesthetized for each MRI session and then put back in their social groups housed at the Station de Primatologie CNRS. For each subject's MRI scan, the PT's surface area was manually traced in both hemispheres using ITK-SNAP software (see Marie et al., 2018).

We found, for the first time in nonhuman new-born, a human-like significant PT's surface asymmetry in favor of the left hemisphere, a brain asymmetry which increases across ages. This finding in non-linguistic primate infants strongly questions the idea that early PT asymmetry constitutes a robust marker for speech development in humans. It also suggest that such asymmetric brain organisation might be (1) highly heritable with a strong genetic component in its development, at both ontogenetic and phylogenetic levels and (2) be inherited from our common ancestor shared with old-world monkeys at least 25-35 million years ago.

Acknowledgements

The project has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programm grant agreement N° 716931 - GESTIMAGE - ERC-2016-STG.

References

Crow, T. (2002). *The speciation of modern Homo sapiens.* Oxford: Oxford University Press.

Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J. F., Le Bihan, D., & Dehaene-Lambertz, G. (2009). Structural asymmetries in the infant language and sensori-motor networks. *Cerebral Cortex, 19*(2), 414-423.

Marie, D., Roth, M., Lacoste, R., Nazarian, B., Bertello, A., Anton, J.-L., Hopkins, W. D., Margioutoudi, K., Love A. S., & Meguerditchian, A. (2018). Left Brain Asymmetry of the Planum Temporale in a non-Hominid primate: Redefining the origin of brain specialization for language. *Cerebral Cortex, 28,* 1808-1815.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29(2),* 143-178.

Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage, 30,* 1414-1432.

OLDER ENGLISH WORDS ARE MORE POLYSEMOUS

ALEKSANDRS BERDICEVSKIS

aleksandrs.berdicevskis@gu.se
Språkbanken (The Swedish Language Bank), University of Gothenburg
Gothenburg, Sweden

Word meanings change over time, usually following routes predicted by general cognitive principles. While significant advances in understanding lexical meaning change have been made, relatively few studies have focused on large-scale quantitative testing of the proposed meaning change laws. It has, for instance, been hypothesized that older words are on average more polysemous, since they have had more time to develop new meanings through meaning shifts. I perform a large-scale quantitative test of this hypothesis, extracting data for 16K English verbs, 45K adjectives and 102K nouns from the Oxford English Dictionary. I show that the hypothesis holds, but the correlation between age and polysemy depends on the word frequency, being stronger for the more frequent words.

1. Introduction

Studying semantic change can give us insights into language evolution, if we manage to understand cognitive processes that underly the change better (Hoefler, & Smith, 2008). An important type of semantic change is change in lexical meaning. Word meanings vary and change, usually following routes predicted by general cognitive principles, such as metaphor, metonymy, generalization and specialization (Nerlich, & Clarke, 2003).

While significant advances in understanding lexical meaning change have been made, relatively few studies have focused on large-scale quantitative testing of the proposed meaning change laws (but see, for example, Urban, 2011; Hamilton, Leskovec, & Jurafsky, 2016; Winter, Thompson, & Urban, 2014; Xu, Malt, & Srinivasan, 2017).

In this paper, I quantitatively test the assumption that older words are on average more polysemous (Lee, 1990). Since words become polysemous through meaning shifts, it is reasonable to expect that older words, which have had more time to develop additional meanings, would have done so.

While a plausible hypothesis, this is not necessarily true. Other factors might dwarf word's age and/or interact with it in complicated ways. Besides, meanings not only emerge, but also disappear, and, applying the same logic, one can predict that older words have had more chances to lose the existing meanings. Since the rates of the emergence and disappearance of lexical meanings are unknown, we cannot claim with certainty which of these diachronic process is dominant.

In other words, whether older words are more polysemous is an empirical question. I am aware of but two studies that address it empirically. Lee (1990) demonstrated that word age positively correlates with polysemy for two samples of 200 English nouns and one sample of 208 English adjectives. Flieller and Tournois (1994) studied a sample of 998 French nouns, and while, having other research questions, they did not focus on the relation between age and polysemy, they also report a positive correlation.

In this paper, I demonstrate positive correlation between word age and polysemy for three parts of speech (verbs, adjectives and nouns), not restricting myself to small samples, but using all words available in the Oxford English Dictionary (OED Online, 2019). The correlation coefficients I report can be used to quantify the average rate at which words develop new meanings.

2. Materials and methods

I browse the online edition of the OED,[1] extracting for every word its part of speech, number of separate meanings, date of entry and frequency.

I focus on three parts of speech: verbs, adjectives and nouns. Parts of speech may differ notably in their semantic behavior (and how lexicographers analyze its behavior), which is why I perform all comparisons only within parts of speech. For technical reasons, I ignore entries that ascribe two different parts of speech to a single lemma (e.g. *Aalenian, n.* and *adj.*). This, however, happens rarely: in most cases, if a word is polysemous across parts of speech, then each part of speech has its own entry (e.g. *iron* has separate entries as a noun, an adjective and a verb).

Homonyms (i.e. words that have the same graphical form, but are assigned to different entries, e.g. *abate*$_1$ ‘to end’ and *abate*$_2$ ‘to seize’) are treated as different words.

Entries marked as obsolete (by a cross † preceding the headword) are ignored.

For most entries, the OED provides the year when the word has first been attested in writing. While this, of course, is just an approximation to the real age of the word, it is as good as we can hope to get. Entries where no date is provided

---

[1] http://www.oed.com/, accessed April 2019

or where the information is considered unreliable (preceded by *ca* or *ante*, or represented as e.g. *17..*) are ignored, as are entries where *OE*, *ME* (resp. Old English, Middle English) etc. is provided instead of year. For date ranges like *1641-1642*, the year before the hyphen is treated as the date of entry. For early periods, the OED does not provide exact years (using instead notation like *OE*). However, automatic browsing results in small number of entries with suspiciously early dates (e.g. *170* or *688*). Manual check shows that most, if not all, entries with years earlier than 951 are due to errors at the OED website. For this reason, they are also ignored.

In order to establish how polysemous a word is I calculate a number of meanings listed within the entry. The OED distinguishes meanings at several levels: overarching meanings (labelled by Roman numerals), more specific meanings within each Roman-numeral meaning (labelled by Arabic numerals), submeanings within each Arabic-numeral meanings (labelled by small letters). I count the Arabic-numeral meanings, since they are closest to most traditional understandings of "different meanings of the same word". If there are no Arabic numerals within the entry, the word is considered to have a single meaning. Obsolete meanings, marked by a cross before the Arabic numeral, are ignored. If there is no cross, the meaning is *not* considered obsolete (and thus included in the analysis), even though there might be a note like *obsolete* or *archaic* within the definition. The reason is that the positioning and wording of such notes is not systematic and they cannot be reliably processed automatically. If all the Arabic-numeral meanings within the entry are obsolete, the word is ignored.

Frequency has been shown to be a major factor affecting polysemy (Hernández-Fernández et al., 2016; Fenk-Oczlon, & Fenk, 2010; Zipf, 1945). The OED entries do not contain exact frequency data, but they provide a frequency band the word belongs to, ranging from 1 (extremely rare) to 8 (very frequent).

It would have been better to use continuous frequency data rather than binned, but in order to obtain accurate frequency estimates a substantial amount of manual work is required (dealing with spelling variation, homonyms, morphological forms; comparing data from different corpora). Since this work has already been done by the OED editors when estimating frequency bands, I am relying on their data.

Some OED entries may differ from the principles that the automatic extraction described in this section relies upon, either due to different editorial policies in different periods or random errors and inconsistencies. This means there might be some noise in the data. Some entries containing obvious mistakes were manually removed, and spot checks did not reveal neither systematic biases nor random errors.

See supplementary materials for the scripts for processing the OED entries, the extracted data and the scripts for statistical analysis.

3. Results

Distribution of word counts per frequency bands and parts of speech is summarized in Table 1.[2]

Table 1. Distribution of word counts per frequency bands and parts of speech.

	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7	In total
Adjectives	5420	22998	9146	5345	1609	182	4	44704
Nouns	11013	43789	26922	15036	4555	883	72	102337
Verbs	1317	6317	3957	2961	1207	307	35	16101

For illustration purposes, the relation between age and polysemy for nouns from band 6 is represented on Figure 1.

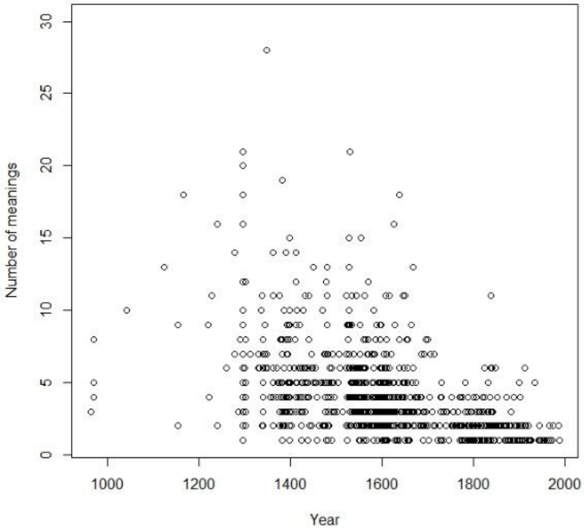


Figure 1. Number of meanings and year of entry for all nouns from frequency band 6.

---

[2] Interestingly, the distribution does not entirely follow the one that could be expected according to Zipf's law: there are always more words in band 2 than in band 1. It is probably explained by the fact that very infrequent words are less likely to get into a dictionary.

To estimate the effect of word age, I fit a Poisson regression model with number of meanings as the dependent variable, date of entry (YEAR) as a continuous predictor, part of speech (POS) as a categorical predictor (with adjectives as the reference level) and FREQUENCY BAND as a reverse-Helmert-coded categorical predictor. All two-way and three-way interactions are also included. To make the intercept more interpretable, it was set to the year 950 instead of 0 (the earliest words included in the analysis are dated 951).

The summary of the regression model are presented in Table 2. For brevity's sake, only the coefficients for YEAR, FREQUENCY BAND, POS and the two-way interactions between YEAR and the other two predictors are included (see the supplementary materials for the full summary of the model).

Table 2. Summary of the Poisson regression model: polysemy as predicted by year of entry, frequency and part of speech. Asterisk (*) marks significance at 0.05 level. See main text for more details.

Coefficient	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**
(Intercept)	1.4e+00	1.1e-01	12.4	<0.001*
year	-1.0e-03	2.1e-04	-4.9	<0.001*
freq.band 2	1.9e-01	5.9e-02	3.2	0.002*
freq.band 3	2.3e-01	2.6e-02	8.7	<0.001*
freq.band 4	2.3e-01	1.7e-02	13.4	<0.001*
freq.band 5	2.1e-01	1.8e-02	11.9	<0.001*
freq.band 6	1.8e-01	3.5e-02	5.3	<0.001*
freq.band 7	2.7e-01	1.0e-01	2.6	0.008*
POS noun	1.7e-01	1.1e-01	1.5	0.130
POS verb	4.2e-01	1.2e-01	3.4	<0.001*
year × freq.band 2	-1.7e-04	6.8e-05	-2.5	0.013*
year × freq.band 3	-2.0e-04	3.1e-05	-6.5	<0.001*
year × freq.band 4	-2.0e-04	2.2e-05	-9.3	<0.001*
year × freq.band 5	-1.7e-04	2.5e-05	-6.9	<0.001*
year × freq.band 6	-1.1e-04	5.1e-05	-2.2	0.030*
year × freq.band 7	-1.1e-04	2.0e-04	-0.6	0.570
year × POS noun	-1.6e-04	2.1e-04	-0.8	0.451
year × POS verb	-3.6e-04	2.3e-04	-1.6	0.119

YEAR has a negative coefficient which is significantly different from zero, which means that older words do indeed have more meanings. FREQUENCY BANDS always have positive coefficients (reverse Helmert coding means that we are comparing words from band 2 with words from band 1, words from band 3 with words from bands 2 and 1, etc.). This reflects the well-established fact that more frequent words tend to be more polysemous (Hernández-Fernández et al., 2016; Fenk-Oczlon, & Fenk, 2010; Zipf, 1945). Verbs, according to the model, are significantly more polysemous than adjectives, while nouns are not.

All but one interactions between YEAR and FREQUENCY BAND have significant (but small) negative coefficients, which means the negative slope is steeper for higher bands. The only exception is band 7, probably due to the very small number of words in it. In other words, for more frequent words age matters more in terms of polysemy; the difference between older and newer words is larger. Interestingly, Lee (1990) does not observe an interaction effect between frequency and age in his data.

Among the coefficients that are not listed in Table 2, five are significant: the interaction between FREQUENCY BANDS 4, 5, 6 and POS verb (0.16, 0.06 and 0.09 respectively), between FREQUENCY BAND 5 and POS noun (0.05), between YEAR, FREQUENCY BAND 4 and POS verb (-8.7e-05); see supplementary materials for further details.

4. Discussion

One goal of the computational approaches to semantic change is to discover fundamental patterns of meaning evolution. Hamilton, Leskovec and Jurafsky (2016), for instance, provide evidence for *the law of conformity* (more frequent words have slower rate of semantic change) and *the law of innovation* (more polysemous words have higher rate of semantic change). This paper provides evidence for *the law of age*: older words are more polysemous.

The estimated rates of change, reported in Table 2, vary across parts of speech and words of different frequency. Apart from confirming that more frequent words are more polysemous, the results show that words from higher frequency bands develop new meanings at faster rates than words from lower bands, i.e. that the correlation between age and polysemy is stronger for frequent words. More detailed investigation using continuous frequency data would be required to understand the interaction between age and polysemy more precisely.

Quantification of semantic change enables us to test the existing qualitative theories about meaning: do the observed results fit with the theoretical predictions? Quantification also makes it possible to predict future changes or to reconstruct the earlier stages of the language.

Further research avenues can include:

- reproducing the study using corpus data instead of dictionary data (to estimate both the age of the word and the number of meanings, using automated sense-induction methods), although that would require large high-quality diachronic corpora;
- reproducing the study for other languages;
- quantifying the rate of disappearance of existing meanings;

- collecting more data about when new meanings appear (the year of the first known usage is provided in the OED for every meaning) in order to explore whether the trajectory is linear or has some other form;
- establishing semantic relations between older and newer meanings (is the new meaning the result of a metaphorical shift, or bleaching, or something else?). That would require either extensive manual annotation or high-quality automatic tools.

Supplementary materials

See https://github.com/AleksandrsBerdicevskis/polysemy.

References

Fenk-Oczlon, G., & Fenk, A. (2010). Frequency effects on the emergence of polysemy and homophony. *International Journal of Information Technologies and Knowledge, 4(2),* 103–109.

Flieller, A., & Tournois, J. (1994). Imagery value, subjective and objective frequency, date of entry into the language, and degree of polysemy in a sample of 998 French words. *International Journal of Psychology, 29(4),* 471–509.

Hamilton, W., Leskovec, J., & Jurafsky, D. (2016). Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).* Association for Computational Linguistics, pp. 1489–1501.

Hernández-Fernández, A., Casas, B., Ferrer-i-Cancho, R., & Baixeries, J. (2016). Testing the Robustness of Laws of Polysemy and Brevity Versus Frequency. In: Pavel Král, Carlos Martín-Vide (eds). *Statistical Language and Speech Processing. SLSP 2016. Lecture Notes in Computer Science*, 9918, 19–29. Springer, Cham.

Hoefler, S., & Smith, A. D. (2008). Reanalysis vs. metaphor? What grammaticalisation can tell us about language evolution. *In The evolution of language: Proceedings of the 7th International Conference (EVOLANG7),* pp. 163–170.

Lee, Christopher. (1990). Some hypotheses concerning the evolution of polysemous words. *Journal of Psycholinguistic Research, 19, 4,* 211–219.

Nerlich, B., & Clarke, D. (2003). Polysemy and flexibility: introduction and overview. In Brigitte Nerlich, David D. Clarke (eds). *Polysemy. Flexible Patterns of Meaning in Mind and Language*. Mouton de Gruyter, pp. 3–30.

OED Online. (2019). *Oxford English Dictionary Online.* Oxford University Press. Accessed April 2019.

Urban, M. (2011). Asymmetries in overt marking and directionality in semantic change. *Journal of Historical Linguistics*, 1, 3–47.

Winter, B., Thompson, G., & Urban, M. (2014). Cognitive factors motivating the evolution of word meanings: Evidence from corpora, behavioral data and encyclopedic network structure. In *Evolution of Language: Proceedings of the 10th International Conference (EVOLANG10)*, pp. 353–360.

Xu, Y., Malt, B. C., & Srinivasan, M. (2017). Evolution of word meanings through metaphorical mapping: Systematicity over the past millennium. *Cognitive Psychology*, 96, 41–53.

Zipf, G. (1945). The meaning-frequency relationship of words. *The Journal of General Psychology*, *33(2)*, 251–256.

PRIMATE PROTO-VOWELS AND THE EARLY EMERGENCE OF SPEECH

LOUIS-JEAN BOË[1] and THOMAS R. SAWALLIS[*2]

[*]Corresponding Author: tom.sawallis@google.com
[1] GIPSA-lab, CNRS, Grenoble INP & U. Grenoble Alpes, Grenoble, France
[2]New College, U. of Alabama, Tuscaloosa AL, USA

1. Laryngeal Descent Theory Refuted

The publications elaborating the Laryngeal Descent Theory (LDT) began 50 years ago (Lieberman, Klatt, & Wilson, 1969). LDT eventually claimed that only anatomically modern *Homo sapiens* (AMHS) could have full language, because only they had the vocal tract (VT) anatomy required to produce contrasting vowels. Other primates, Neanderthals included, were limited to the schwa-like vocalizations generated by a uniform-tube configuration of the VT.

Criticism of LDT began about 20 years ago, and culminated recently in two technologically updated replications of early LDT studies. Fitch et al. (2016) used x-ray videos of macaque facial gestures to estimate the VT's area function and then synthesized the vowels implied by those configurations. Boë et al. (2017) analyzed naturally produced calls by Guinea baboons. In normalized F1-F2 space, the synthetic macaque vowels exceeded LDT's predicted schwa, covering about a quarter of the vowel triangle, and included tokens in /u/ and /ɑ/ zones, while the baboon tokens occupy over twice the macaque space, and include tokens in /ɨ æ ɑ ɔ u/ zones. Both species lack key elements of AMHS vocal anatomy, so their contrasting vowel qualities refute LDT's claim that only AMHS can produce non-schwa vowel qualities.

2. Contrasting Proto-Vowels Discovered in Previous Primate Studies

Extending the search for primate vowel contrasts requires tools. We review how vowel spaces are normalized for cross-species comparison using VT length

(VTL), and how to estimate VTL from formant values when anatomic measurements are unavailable. We also document functions for schwa formant frequencies by VTL and for certain known primate VTLs. We then examine a selection of articles (1993 – 2016) for potential non-schwa vowels in other living primates. Our analysis (in press, *Science Advances*, Dec. 2019) reveals vowel qualities outside the schwa region in gorillas, baboons, macaques, Diana monkeys, and even lemurs. We conclude that the ability to produce potentially contrastive non-schwa vowel qualities dates back at least to our last common ancestor with Cercopithecoidea (Old World monkeys) some 27 Ma ago.

3. The Dawn of Speech vs. the Dawn of Language

The LDT was early and influential in the burgeoning field of language evolution, and has been cited since by researchers (up through, e.g., Hauser et al., 2014) as reinforcing the case for language emergence as recently as 100 – 70 ka ago. Their scenario is a recent, sudden, and simultaneous emergence of speech and language triggered by the speciation event distinguishing AMHS from previous hominid forms, about 300 ka ago.

Our findings render that scenario untenable. The Dawn of Speech in the form of contrasting vowel sounds is not recent, but early, about 100x earlier than AMHS speciation. Speech emergence was therefore not sudden, but extended, probably via stages now inviting both theory and investigation. Its final developments necessarily coincided with language emergence, but the duration difference, 2 orders of magnitude, makes simultaneity meaningless. Language, whenever it emerged, had to integrate into some previously developed speech system.

References

Boë, L.-J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., … Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PLOS ONE*, *12*(1), e0169321. https://doi.org/10.1371/journal.pone.0169321

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, *2*(12), e1600723.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., Chomsky, N., Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, *5*, 401.

Lieberman, P., Klatt, D. H., & Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, *164*(3884), 1185–1187.

THE ADVANTAGE OF EXTREME MEANINGS IN CULTURAL EVOLUTION

FAUSTO CARCASSI*[1], MARIEKE SCHOUWSTRA[1], and SIMON KIRBY[1]

*Corresponding Author: fausto.carcassi@gmail.com
[1]Centre for Language Evolution, University of Edinburgh, UK

Gradable adjectives are used to convey that an individual has a property to a degree greater (or to a degree lower) than some threshold. For instance, “Mary is tall” conveys that Mary’s height is greater than some threshold on the scale of heights. These thresholds can in principle be anywhere, but a remarkable connection has been shown between the thresholds’ positions and the structure of the underlying scale. Namely, if the scale is bounded adjectives tend to be at the maximum or minimum of the scale. This tendency is surprising because generally, communication is more accurate when categories have roughly the same size (Jäger, Metzger, & Riedel, 2011). Various models for adjectival semantics attempt to explain the boundary-minded behaviour of adjectives like ‘full’ by appeal to online pragmatic considerations (Lassiter & Goodman, 2015), long-term communicative efficiency (Qing & Franke, 2014), the saliency of extreme points and its role in referential language use (Kennedy, 2007; Potts, 2008; Franke, 2012).

We present an account of extreme thresholds, defined as thresholds positioned at a scale’s border, that does not make assumptions about specific cognitive preferences, but rather only appeals to the mechanisms of learning. We claim that extreme thresholds have an advantage in cultural evolution because they produce data that can be learned with a greater accuracy than data produced by non-extreme thresholds. Once a population stumbles upon an extreme language, the language is transmitted with high fidelity and tends to persist in the population. We support this explanation with a computational model that combines Iterated Learning (IL) as a model of cultural evolution (Kirby, Griffiths, & Smith, 2014; Kirby, Tamariz, Cornish, & Smith, 2015) and Rational Speech Acts (RSA) as a model of pragmatic communication (Goodman & Frank, 2016).

The language in the model consists of three signals: s_σ (silence), s_+ (positive polarity adjective like “full”), and s_- (negative polarity adjective like “empty”). Each signal conveys that a degree d_o falls in a certain part of the relevant scale. s_σ is compatible with the whole scale, leaving the position of d_o unspecified. s_+ [s_-] conveys that d_o is greater [lower] on the scale than a value θ_+ [θ_-]. Agents are pragmatic speakers that, given an observation, tend to produce the signal s that is most useful for the listener. The utility of s is calculated as the expected distance between the listener's guess and the speaker's observation, given s. This causes the speakers to allow *pragmatic slack* (Lasersohn, 1999), i.e. speakers sometimes use signals that are not literally compatible with the observation. Pragmatic slack is regulated by a λ parameter for each signal; the greater the λ parameter, the less pragmatic slack is allowed. The λ parameter is different from the RSA rationality parameter, which is fixed at 4. In the model, we always set the λ parameters for s_+ and s_- to the same value.

The IL model consists of chains of single agents. The agent in the i^{th} position in the chain, a_i, produces linguistic data D_i consisting of tuples of observations and signals. a_{i+1} performs Bayesian inference on D_i to learn a_i's θ_+ and θ_-. The learner's prior over θ_+ and θ_- is uniform. Since the first agent in the chain has no cultural parent, a_0's θ_+ and θ_- are picked randomly. We simulate chains of 10000 agents for various combinations of λ parameters and numbers of observations.

Fig 1 shows the proportion of extreme thresholds that evolve for each combination of parameters. Uniformly sampled thresholds would (almost surely) not be extreme. However, we observe that, for some combinations of parameters, up to a quarter of the evolved thresholds are extreme. This shows that extreme thresholds have an advantage in cultural evolution, that increases when there are fewer observations and when there is more pragmatic slack.

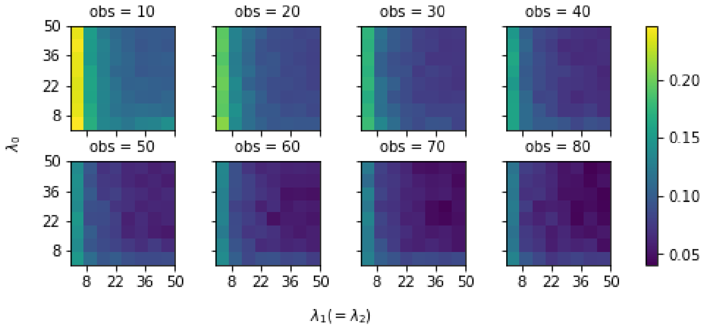


Figure 1. Results of IL for various parameters values. Obs is the number of observations made by learners. λ_σ, λ_+, and λ_- regulate the pragmatic slack of s_σ, s_+, and s_- respectively. Color indicates the proportion of all meanings that are extreme, where extreme means 1. for $\theta_+$and 0. for θ_-.

In sum, we show that extreme thresholds have an advantage in cultural evolution that does not come from a prior preference for extremeness, but rather from the type of data that extreme thresholds produce. Moreover, we show that this advantage is modulated by the amount of data that learners observe and the acceptable level of pragmatic slack.

References

Franke, M. (2012). On scales, salience and referential language use. In *Logic, Language and Meaning* (pp. 311–320). Springer.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Jäger, G., Metzger, L. P., & Riedel, F. (2011). Voronoi languages: Equilibria in cheap-talk games with high-dimensional types and few signals. *Games and Economic Behavior*, *73*(2), 517–537.

Kennedy, C. (2007). Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, *30*(1), 1–45.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Lasersohn, P. (1999). Pragmatic Halos. *Language*, *75*(3), 522–551.

Lassiter, D., & Goodman, N. D. (2015). Adjectival vagueness in a Bayesian model of interpretation. *Synthese*, *194*(10), 3801–3836.

Potts, C. (2008). Interpretive economy, Schelling points, and evolutionary stability. *Manuscript, UMass Amherst.*

Qing, C., & Franke, M. (2014). Gradable adjectives, vagueness, and optimal language use: A speaker-oriented model. In *Semantics and linguistic theory* (Vol. 24, pp. 23–41). New York.

DE-CODING THE EVOLUTION OF COMMUNICATION: HOW THE CODE MODEL SYSTEMATICALLY BIASES STUDIES OF ANIMAL COMMUNICATION

ERICA A CARTMILL

Departments of Anthropology and Psychology, University of California, Los Angeles, Lox Angeles, USA

Humans have highly-developed ostensive and inferential capacities, such that virtually any behavior can be used as a communicative signal (Sperber & Wilson, 1995; Origgi and Sperber 2000; Scott-Phillips, 2014). Some authors propose that these capacities make human language different *in kind* from other primate communication systems and, indeed, from animal communication more broadly (e.g., Scott-Phillips, 2014, 2015). According to Scott-Phillips, primate communication relies on “natural codes” (ones that derive their meaning largely through inheritance), while human language relies on “conventional codes” (ones that derive their meaning through learning and convention). However, this characterization of human language as infinitely flexible and expandable and primate communication as largely inflexible and fixed does not take into account the great difference in methodological approaches to studying primate communication and human language.

Any primatologist will tell you that primates interact in many complex, difficult-to-measure ways. The trick is trying to capture that complexity in an operationalizable, repeatable way so that it can move from anecdote to data. Animal communication research is built around searching for replicable contingencies between a signal given by one individual and a subsequent change of behavior in another. Or, for signals that reference things in the world (like predators), a contingency first between an eliciting context and a signal and then between that signal and the response. This search for animal signals is strongly shaped by the “code model”, which grew out of the Shannon model of information (Shannon 1948) and the conduit model of information transfer (Reddy 1979). This metaphor has had a lasting effect on research design, exaggerating the gap between human language and primate communication.

Studies of primate communication find codes. But this is unsurprising, since studies of primate communication are *looking for* codes. To put it another way, the conclusions are largely built in to the methods. In this talk, I use studies of ape gesture and vocalization as example cases to highlight the ways in which characterizations of ape communication systems are strongly influenced by the underlying assumptions of the code model, which influence data collection and analysis. While studies are designed to identify signal-meaning pairs, selective data practices and publishing biases also contribute to characterizing primate communication as codes. For example, in ape gesture studies, datasets are typically narrowed down considerably to include those that are likely to show predictable associations.

It is unsurprising that animal communication literature focuses on these predictable elements. They are the most replicable, they can demonstrate that a signal has meaning, and they can reveal some of the properties of human language (i.e., reference). The assumptions of this model are insidious. In a recent article co-authored by some of the biggest names in animal communication, the authors urge the field to adopt a standardized definition of intentional communication that avoids mental-state attribution. But even this definition revolves around observing code-like contingencies in signals and responses (Townsend et al., 2017). It may be impossible (or even undesirable) to eschew searching for clear contingencies between signals and responses, but we should be cautious when concluding that a species' communication system is primarily a code when that is what we have looked for.

Comparative studies across primate species have great potential to reveal the homologous features of communication and cognition that gave rise to language in the human lineage. But, in seeking to characterize the differences between human and non-human communication, it is critical that we do focus our search too strongly on identifying simple codes, lest we discover only those. The theory that human communication is built on a framework of ostension and inference is compelling, but to determine whether humans are unique in these abilities, we must assess the lasting impact of the code model framework on studies of animal communication and ensure that our conclusions are not driven by differences in our methods.

References

Origgi, G., and D. Sperber. (2000). Evolution, communication and the proper function of language. In *Evolution and the human mind: language, modularity and social cognition.* P. Carruthers and A. Chamberlain, eds. Cambridge: Cambridge University Press. Pp. 140–169.

Reddy, M. J. 1979. The conduit metaphor: a case of frame conflict in our language about language. In *Metaphor and thought.* A. Ortony, ed. pp. 284– 324. Cambridge: Cambridge University Press.

Scott-Phillips, T. (2014). *Speaking Our Minds: Why human communication is different, and how language evolved to make it special.* Palgrave MacMillan.

Scott-Phillips, T. (2015). Nonhuman primate communication, pragmatics, and the origins of language. *Current Anthropology*, *56*(1), 56-80.

Shannon, C. E. 1948. A mathematical theory of communication. *Bell Systems Technical Journal* 27:379–423.

Sperber, D., and D. Wilson. (1995). *Relevance: communication and cognition.* 2nd edition. Oxford: Blackwell.

Townsend, S.W., Koski, S.E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., Glock, H.J., Goncalves, I.B., Burkart, J.M., Flower, T., Gaunet, F., Glock, H.J., Gruber, T., Jansen, D., Liebal, K., Linke, A., Miklósi, A., Moore, R., van Schaik, C., Stoll, S., Vail, A., Waller, B.M., Wild, M., Zuberbühler, K., and Manser, M.B. (2017). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biological Reviews*, *92*(3), 1427-1433.

DID POST-NEOLITHIC CHANGES IN BITE CONFIGURATION IMPACT SPEECH? A NEW APPROACH TO THE QUESTION

SIHAN CHEN [1], CALEB EVERETT* [2]

*Corresponding Author: caleb@miami.edu
[1] Department of Mechanical and Aerospace Engineering, University of Miami, USA
[2] Department of Anthropology, University of Miami, USA

Blasi et al. (2019) offer evidence that post-neolithic changes in bite configuration, owed to the adoption of agriculture, have led to the innovation and proliferation of labiodental consonants in the world's languages. Here we investigate the putative association between agriculture and labiodental consonants via a new approach that does not rely on phoneme inventories. Given that labiodentals are apparently characterized by reduced muscular effort in populations with agriculture-influenced bite configurations, we test whether labiodental sounds are actually more prevalent in languages whose speakers rely on agriculture. We rely on word lists from the Automated Similarity Judgement Program (Wichmann et al. 2018), which contains transcribed lists of common words in thousands of languages. We analyze the relative frequency of sound types in the word lists of agricultural and hunter-gatherer populations, respectively, finding differing mean rates of labiodental usage in populations with distinct subsistence strategies. Using a linear mixed-effects model to control for relatedness and contact, we find support for an association between the frequency of labiodental consonants and the use of agriculture.

1. Introduction

Do minor cross-population variations in vocal tract anatomy foster disparities in the sound systems used in languages? The "uniformitarian hypothesis" prevalent in linguistics maintains that languages evolve in ways that are not significantly impacted by such variations, yet this hypothesis has been called into question by recent research on several sound types. (Dediu & Moisik 2019, Dediu, Janssen, & Moisik 2019, *inter alia*) Most prominently, perhaps, Blasi et al. (2019) suggest that post-neolithic changes in the bite configurations of agricultural populations have yielded effects on the phoneme inventories of the world's languages: Labiodental consonants are now relatively common in the world's languages, putatively due to the reduction of bite-to-bite configurations and the increase in overjet and overbite owed to the softer foods characterizing agricultural diets.

Blasi et al. support this intriguing hypothesis, first suggested by Hockett (1985), with a series of findings. Perhaps most crucially, those findings include a worldwide association between labiodental consonants and agriculture judging from the roughly 2,000 cultures considered. This association is crucial to the hypothesis under question since, were it absent we would have little evidence to believe that the modeled reduced muscular effort, characterizing labiodentals in mouths with overbite and overjet (when contrasted to "flat" bites), is sufficient to meaningfully impact speech. After all, it is known that there are minor cross-population differences in vocal-tract anatomy. The question is whether such differences are actually sufficient to have any meaningful effect on speech. While Blasi et al. (2019) offer compelling diachronic evidence for their case, such historical evidence is based primarily on Indo-European languages. In short, the worldwide distribution of labiodental sounds is essential to illuminating this issue. Here we investigate this distribution with a new and complementary method. We aim to contribute not just to the specific hypothesis promoted by Blasi et al. (2019), but also to the larger question of whether extra-linguistic factors influence the ways that languages evolve by creating selective pressures for/against some features. Such factors have been suggested in other recent work. (e.g. Everett 2017)

2. New approach

Blasi et al. (2019:6) suggest that "labiodental production effort" is reduced, by about 30% in fact, in populations without the edge-to-edge bite configuration that is characteristic of hunter-gatherer groups. Given that articulatory effort is at the heart of the tested hypothesis, one could argue that the phoneme data on which Blasi et al. rely offer an essential but still incomplete depiction of the relevant typological data. If the trend towards the inclusion of labiodental sounds in a language is (partially) the byproduct of articulatory ease and production-effort reduction, then we might expect that speakers with edge-to-edge bite generally rely on labiodental sounds less in speech. Conversely, we might expect that speakers with overbite and overjet rely on such sounds more in speech. Arguably, phonemic status is a proxy for what we are ultimately interested in, viz. the rate of occurrence of labiodental sounds in the speech stream. Consider the following points, which underscore the need for examining the rates of occurrence of a sound as opposed to only examining its binary phonemic status in a language: A language may have a given phoneme, but that phoneme may be rare in speech—perhaps contrastive in only a few minimal pairs. If a "rare" labiodental phoneme exists in the language of a group of agriculturalists, this case would support Hockett's hypothesis under a phoneme-based analysis. But it would arguably not support the hypothesis nearly as well as a case in which a labiodental phoneme was frequent in the speech of a culture of agriculturalists. The converse possibility also underscores the desirability of the approach we pursue below: Suppose a language of hunter gatherers has no labiodental phonemes, but labiodentals do occur in speech as allophones of some related phoneme. For example, perhaps a phonemic bilabial stop is lenited word-finally. Such a scenario would be considered consistent with the hypothesis under Blasi et al.'s analysis, but would be dissatisfying from the perspective of sound usage. After all, the phonetic realization of a sound is what actually requires muscular effort, and the reduction of muscular effort can only be realized in phonetic patterns. While the biomechanical modeling in Blasi et al. (2019) clearly suggests that labiodental consonants require less production effort in populations without edge-to-edge bite, to test whether this reduction actually impacts speech significantly we should also consider how phonemes are reified phonetically. To be clear, we are not arguing against the phonemic approach utilized in Blasi et al. (2019), and in fact the consideration of phoneme inventories is also critical, particularly as it sheds light on the diachronic claim central to Hockett's original hypothesis. Yet the rate of occurrence of relevant phonetic units in speech is, in our estimation, another factor to consider in testing the hypothesis.

The suggestion that all phonemes do not equitably represent phonetic patterns in speech is supportable with specific examples. Consider, for instance, the voiced postalveolar fricative (/ʒ/). This sound is phonemic in English but only because of a few minimal pairs (e.g. "beige" [beʲʒ] vs. "base" [beʲs]). Recent analysis of the frequency of English phonemes has found that this consonant represents about 0.2% of sounds in large corpora of speech. The most common consonantal phoneme in those same data, /n/, represents about 13% of all sounds. So one consonantal phoneme is about 63 times as common as the other, meaning they are not equally representative of phonetic patterns in a language. (Chin et al. 2012) A recent meta-analysis of studies on 32 languages' sound systems observed that a sound's frequency in phoneme inventories across languages is not always a good indicator of its frequency in actual speech within languages. (Gordon 2016) Some sounds are less frequent in speech than we might expect given their commonality in phoneme inventories. This was found to be true with respect to the voiceless labiodental fricative /f/, the most common sort of labiodental phoneme. (Gordon 2016)

3. Methods & Results

We examined the largest database of phonetically transcribed word lists, the Automated Similarity Judgment Program. This database contains between 40-100 words for each of about 7000 language varieties. This lends typological breadth to our approach, though with clearly limited depth. This limitation is being addressed in follow-up work with other data sources, though it should be noted that the 40-100 common words in the ASJP data are generally frequent in speech and are often reasonable indicators of more pervasive sound patterns in the represented languages. (Everett 2018) The languages in the ASJP database were cross-referenced with the same subsistence database used by Blasi et al. (2019) (derived from Güldemann et al. (2019)), allowing us to rely on the same principal subsistence categorization of languages/cultures. This approach yielded 2756 data points for which word list data could be contrasted with the subsistence-strategy data. For each of the associated 2756 word lists, the total number of labiodental tokens was tabulated. This total was then divided by the entire number of consonant tokens, for each word list (vowels and non-segmental symbols were ignored). This yielded a "labiodental ratio" value for each of the word lists. We relied on a function written by CE via the *stringr* package in *R*. We also used a function written by SC with MATLAB, and then contrasted the results of SC's approach and CE's approach to ensure that the labiodental ratios obtained were identical. (Data and code available upon request.) The labiodental ratio is the proportion of all consonant segments in a word list that are [f] or [v], as all labiodental sounds in the ASJP data are coded with [f] or [v]. Other labiodental sounds exist but are much rarer. Still, it must be acknowledged that the phonetic transcriptions in the ASJP database are sometimes coarse. Yet the typological breadth they offer creates clear advantages as well.

At the roughest level of analysis, prior to instituting any controls for Galton's problem, Hockett's hypothesis is supported by our approach. As is evident in Figure 1, the dialects of hunter gatherer populations skew towards the lower end of labiodental usage. In Table 1 the mean labiodental ratios of word lists are provided, categorized according to the subsistence strategy associated with the speakers of the dialects represented in the lists. For the 2223 dialects categorized as belonging to non-hunter-gatherers by Güldemann et al. (2019), the proportion of consonantal phonetic segments represented by labiodental consonants averages about 2.2%. In contrast, for the 533 word lists categorized as belonging to hunter-gatherers, the proportion of labiodental consonants is about 0.77%. In other words, labiodental consonants are almost three times as common in languages in agricultural groups, across all word lists in the data.

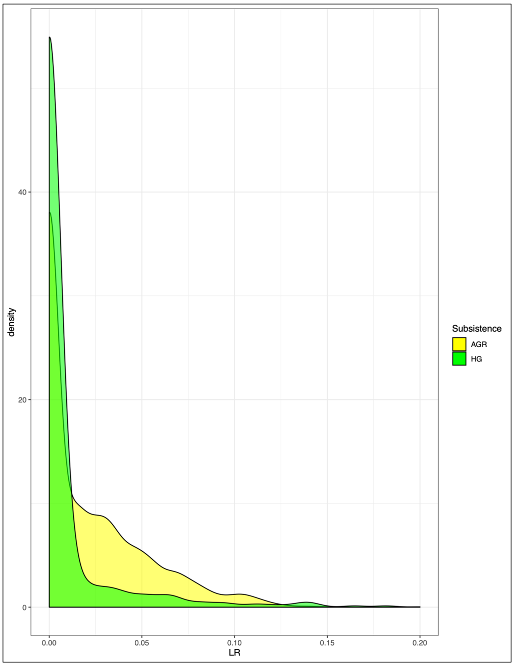


Figure 1. Density distribution of "labiodental ratios", categorized according to subsistence strategy most associated with a dialect's speakers.

Table 1. Proportion of all consonants in phonetically transcribed word lists that are labiodental.

	mean	s.d.
Non-hunter-gatherer (N=2223)	0.0217	0.029
Hunter-gatherer (N=533)	0.0077	0.024

The values in Figure 1 and Table 1 could be distorted by a few language families or linguistic regions, making their interpretation difficult. In Figure 2 the geographic distribution of the top quartile of languages, in terms of prevalence of labiodental consonants, is plotted. These dialects are not randomly distributed geographically. For instance, labiodental consonants are quite prevalent in Europe, though the phylogenetic reconstruction in Blasi et al. (2019) suggests this is a recent phenomenon.

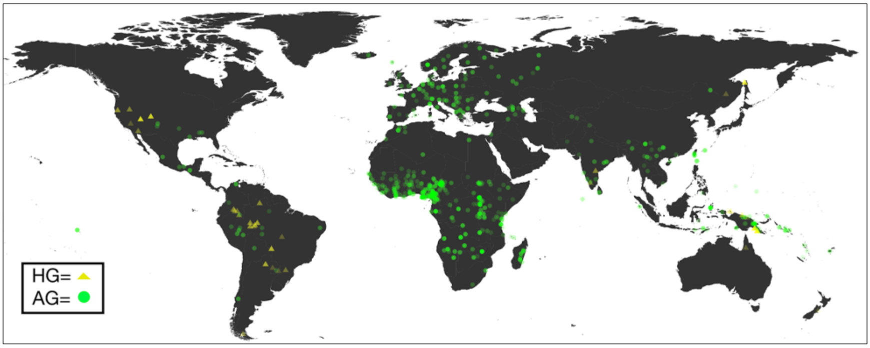


Figure 2. Locations of the languages in the top quartile of labiodental usage, judging from the occurrence of labiodentals in the word lists from the ASJP database. (Increased brightness corresponds to higher labiodental ratios.)

To control for the prevalence of labiodentals in some regions and the effect of large language families whose populations tend to rely on agriculture (e.g. Indo-European), we used a linear mixed effects model. The classifications of word lists into families and geographic regions was based on the AUTOTYP database, which utilizes a relatively fine-grained set of 24 independently motivated geographic regions. This limited the analysis to 1986 ASJP word lists. The *lmer* package in R was utilized with a random-intercepts approach. One model treated subsistence category as a fixed effect, while language family and geographic region were treated as random effects. Since labiodental ratios are technically bounded at 0 and 1, we used logit-transformed LR's as the dependent variable. In a null model, no fixed effect was provided and language family and geographic region were again treated as random effects. A likelihood ratio test contrasting the two models yielded a significant difference. Reliance on hunting and gathering affected the logit-transformed labiodental ratios ($\chi 2$ (1)=11.85, p=0.0006). (Intercept of fixed effect= -3.255, Correlation of fixed effect= -.309.) We interpret this result as strong additional support for Hockett's hypothesis. These results suggest that the clear disparity in labiodental ratios across populations with different subsistence strategies, evident in Table 1, is not simply due to confounds such as language contact and/or relatedness. Nevertheless, the results are based on one linguistic database and one categorization of population subsistence strategies, so they should still be interpreted with caution. In the next section we discuss alternate methods we are utilizing to further investigate this apparent association. Preliminary results with those alternate methods also buttress the accounts of Hockett (1985) and Blasi et al. (2019).

4. Discussion and conclusion

Via a new approach, we have offered evidence that further supports the claims in Hockett (1985) and Blasi et al. (2019). Different sources of data for individual languages are being examined in follow-up work. In particular, we are analyzing texts from the Journal of the International Phonetic Association that also allow for typological breadth. These texts are not Swadesh-type word lists, but short transcribed stories. The results of the analysis of those texts thus far obtained are consistent with those in Table 1. In our follow-up work we are also using an alternate subsistence taxonomy, so that the results are not based too heavily on the classification of cultures in Güldemann et al. (2019). Additionally, the ongoing work focuses not just on labiodental ratios, but on word-initial labiodental ratios. We focus on word-initial sounds given their salience to transcribers and given that they are less likely to be affected by reductive processes. (Wedel et al., In press) We are also examining historical and comparative evidence for the exceptions that have presented themselves in the ASJP data, including a few key Amazonian test cases evident upon careful examination of Figure 2. Finally, our ongoing work includes phonetic analysis of actual individuals who have different bite types. These individuals are speakers of the same language (English), but preliminary work suggests they rely on labiodental consonants to differing degrees.

We believe the results of the present study offer additional support for the hypothesis detailed in Hockett (1985) and carefully followed-up on in Blasi et al. (2019). Labiodental sounds are less frequent in the speech of hunter-gatherers, judging from the rates of occurrence of consonants in common words. Ongoing work is helping to determine whether this pattern holds for other data sets, and whether completely different approaches also lend support to Hockett's intriguing hypothesis.

Acknowledgements

The authors wish to thank three anonymous reviewers for constructive feedback on the original submission. Several changes have been made in accordance with that feedback.

References

Blasi, D., S. Moran, S.R. Moisik, P. Widmer, D. Dediu & B. Bickel. (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* 363.6432

Dediu, D. & S.R. Moisik. (2019). Pushes and pulls from below:Anatomical variation, articulation and Sound Change. *Glossa: A Journal of General Linguistics* 4.7. DOI: http://doi.org/10.5334/gjgl.646

Dediu, D., R. Janssen, & S.R. Moisik. (2019). Weak biases emerging from vocal tract anatomy shape the repeated transmission of vowels. *Nature: Human Behaviour* DOI: 10.1038/s41562-019-0663-x.

Everett, C. 2017. Languages in drier climates use fewer vowels. *Frontiers in Psychology* 8.1285. doi:10.3389/fpsyg.2017.01285

Everett, C. 2018. The global dispreference for posterior voiced obstruents: A quantitative assessment of word list data. *Language* 94.e311-e323.

Gordon, M. 2016. *Phonological typology*. Oxford University Press: Oxford.

Güldemann, T., P. McConvell & R. Rhodes. 2019. The language of hunter-gatherers. Cambridge University Press: Cambridge.

Hockett, C. 1985. Distinguished Lecture. *American Anthropologist* 87:263-281.

Shin, J., J. Kiaer & J. Cha. (2012) *The sounds of Korean*. Cambridge University Press: Cambridge. Chapter 6.

Wedel, A., A. Ussishkin & Adam King. In press. Crosslinguistic evidence for a strong statistical universal: Phonological neutralization targets word-ends over beginnings. *Language.*

Wichmann, S., E. Holman & C. Brown. 2018. The ASJP Database. Available at: https://asjp.clld.org/

TESTING THE LANGUAGE-THROWING HYPOTHESIS

MICHAEL DESMOND COLLINS

Michael Collins: collins.michael@ucdconnect.ie
Psychology Department, Leeds Beckett University, Leeds, UK

1. Introduction

Language and tool use – including object throwing – have long been considered important in human evolution (e.g. Darwin, 1871). Calvin (1982) made a novel proposal for the origin of language capacity in hominins: Morphological adaptations for arboreal living in ancient apes were exapted for object throwing and led to the emergence of hominins as an action-at-a-distance predator. Throwing accurately, however, requires precise timing in the coordination of various effectors, placing a strong selective pressure for multiple morphological changes including more neurons and increased encephalisation. These timing requirements for the sequencing of muscle actions are similar to those required for speech. Calvin hypothesised that the neural mechanisms that supported this timing in throwing were then exapted for use in language. This is the language-throwing hypothesis (Calvin, 1982, 1983, 1993). There is some supporting evidence from comparative research with chimpanzees (Hopkins, Russell & Schaeffer, 2012), but there is a paucity of research in humans. This talk presents a novel experimental method and results testing this idea in humans.

2. Methodology

A Dual-Task (DT) interference paradigm was used in a series of three experiments. The logic of the DT method is that overloading two systems simultaneously creates interference in performance and provides a window into their functional architecture (Pashler, 1994). DT normally uses discrete outcome variables such as hits/misses. However, language and throwing are both highly evolved, massively redundant systems capable of compensating for both natural and experimental perturbations. Thus, performance was analysed not only in terms of *outcome variables* but also in terms of *throwing dynamics*, i.e. how throwers shape their co-ordination. This co-ordination is mapped onto a solution manifold using two outcome variables as axes: velocity and angle of release. Moreover, variability was investigated by applying a Tolerance, Noise and Covariance (TNC) analysis. TNC is a statistical approach to quantify positive and negative variability in co-ordination by comparing actual data with optimal data (Cohen & Sternad, 2009). Finally, as there is a tendency for mouth-movements and hand-movements to synchronise (Vainio, 2019), possible disruption of this tendency was investigated by measuring the difference in milliseconds between maximum stress in vocalisation and the moment of ball release.

3. Experimental Design

In experiment 1, using a repeated measures design, expert throwers (N= 24) threw a tennis ball at a 1.5m target 10m distant while simultaneously vocalizing under four conditions: Complex word, Simple Word, Grunt, Silence (control), using a variation of the Word Complexity Measure (Stoel-Gammon, 2009). Time-stamped video (1000fps) and audio of the throws and vocalizations were recorded for later analysis. Experiment 2 (N=25) measured maximum distance under 2 conditions (Complex word and Silence).

4. Preliminary Results

A linear contrast analysis was conducted on three outcome variables: hits/misses, radial error (i.e. hit distance from the target centre) and maximum distance. A repeated measures ANOVA was conducted on the angular distribution of hits and on two throwing dynamics variables: Release Height and Release Velocity. All found no statistically significant difference ($p > .05$). Finally, a repeated measures ANOVA conducted on the Stress-Release difference (ms) found a small (Partial eta squared $\eta_p^2 = .008$), but statistically significant ($p < .05$) difference between the complex and simple word conditions. This indicates a possible interference effect. However, caution is warranted: multiple DV's were tested in line with the pre-registration statement and the exploratory nature of this research. Therefore, replication is required before assuming this effect is real and not a statistical artefact. Results for TNC analysis are ongoing and will be reported at conference. In conclusion, this novel paradigm found some evidence in favour of the language-throwing hypothesis and technical details (code etc.) will be made available to other researchers.

References

Calvin, W. H. (1982). Did throwing stones shape hominid brain evolution? *Ethology and Sociobiology*, *3*(3), 115–124.

Calvin, W. H. (1983). A stone's throw and its launch window: Timing precision and its implications for language and hominid brains. *Journal of Theoretical Biology*, *104*(1), 121–135.

Calvin, W. H. (1998). The Emergence of Intelligence. *Scientific American*, 7.

Cohen, R. G., & Sternad, D. (2009). Variability in motor learning: Relocating, channeling and reducing noise. *Experimental Brain Research*, *193*(1), 69–83.

Darwin, C. (1981). *The descent of man, and selection in relation to sex*. Princeton, N.J: Princeton University Press.

Hopkins, W. D., Russell, J. L., & Schaeffer, J. A. (2012). The neural and cognitive correlates of aimed throwing in chimpanzees: A magnetic resonance image and behavioural study on a unique form of social tool use. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1585),

Pashler, H. (n.d.). Dual-Task Interference in Simple Tasks: Data & Theory. *Psychological Bulletin*, *116*(2), 220–224.

Stoel-Gammon, C. (2010). The Word Complexity Measure: Description and application to developmental phonology and disorders. *Clinical Linguistics & Phonetics*, *24*(4–5),

Vainio, L. (2019). Connection between movements of mouth and hand: Perspectives on development and evolution of speech. *Neuroscience & Biobehavioral Reviews*, *100*, 211–223.

PHONOLOGICAL CUES TO SEMANTIC CLASS MEMBERSHIP ACROSS HUNDREDS OF LANGUAGES

PABLO CONTRERAS KALLENS[*1] and MORTEN H. CHRISTIANSEN[1]

[*]Corresponding Author: pc684@cornell.edu
[1] Department of Psychology, Cornell University, Ithaca, NY, USA

The cultural evolution perspective suggests that human language is primarily a product of linguistic adaptation to a variety of cognitive, communicative, and social constraints, rather than the result of biological adaptations (e.g., Beckner et al., 2009; Christiansen & Chater, 2008). But how is it possible to acquire complex language without neural mechanisms dedicated to this purpose? One suggestion is that languages, via cultural evolution, “recruit” various types of cues to facilitate learning and use. This implies that all languages should incorporate some constellation of cues (mostly probabilistic in nature) that make them easier to acquire and use (Christiansen, 2013).

Here we empirically explore this multiple-cue approach to language evolution by considering the problem of categorizing words according to their meaning, focusing on the basic distinction between words for actions (typically verbs) and words for objects (typically nouns) (Vigliocco, Vinson, Druks, Barber, & Cappa, 2011). Previous studies have shown that phonological cues, whereby words with similar meanings have some degree of sound similarity (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015), can be observed across a range of languages (Dautriche, Mahowald, Gibson, & Piantadosi, 2017) and can be used for lexical categorization (Monaghan, Christiansen, Farmer, & Fitneva, 2010).

However, most of the previous studies have focused on languages from industrialized and/or literate parts of the world, covering only a few language families. Our analysis uses the word lists available in the Intercontinental Dictionary Series (Key & Comrie, 2015) to show that phonological cues to word class are available in a variety of language families and geographical zones. This database includes phonological transcriptions for more than 200 languages from 59 families, with a median of 911 words per word list. Because the transcriptions for each language in the IDS are not readily comparable, all results are *within* *language*. We used information from Concepticon (List, Greenhill, Rzymski, Schweikhard, & Forkel, 2019) to determine the broad semantic class of a word: action, object or other. We focused on this semantic distinction because it is more fundamental compared to lexical categories such as verbs and nouns. Using a normalized version of the Levenshtein edit distance (Yujian & Bo, 2007), we computed the mean distance to actions and to things for each word by averaging over all within-language pairwise distances. The difference between the former and the latter is then a measure of the a word's *phonological typicality*.

To assess the potential effect of morphological markers, we assigned languages where more than a third of the words in a class share the same final or initial phoneme (167 of the 227 languages) to a "marker" group. A two-sample Wilcoxon rank-sum test shows that the difference in typicality between actions and objects is significant for 162 of the languages with markers and 22 out of the 59 languages without markers, Bonferroni adjusted $p < .001$.

To test whether these phonological cues are useful for learning, we trained an iRNN-type (Le, Jaitly, & Hinton, 2015) recurrent neural network with a 10-unit hidden layer to learn to classify the sequences of phonemes in a word as either an action or a thing. The performance of the network was assessed through Matthews Correlation Coefficient (MCC), where chance is 0 and perfect prediction is 1. Each network's MCC was cross-validated using a 10-fold scheme. Figure 1 shows all the within-language median MCC scores as a function of marker vs no-marker group. A within language one-sided Wilcoxon ranked sum shows that 158 of the 167 languages of the marker group and 27 out of the 45 languages in the no-marker group have an MCC higher than chance.

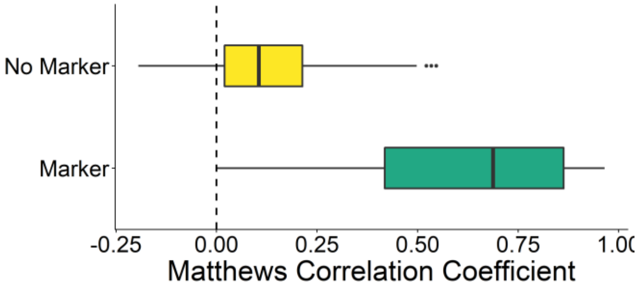


Figure 1 All within-language median MCC scores.

In conclusion, there is strong evidence that a great variety of languages encode broad semantic distinctions in subtle but useful phonological patterns, beyond the potential effects of morphological markers. This provides further evidence that languages have evolved to facilitate key parts of their acquisition.

References

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., … Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning, 59,* 1–26.

Christiansen, M. H. (2013). Language has evolved to depend on multiple-cue integration. In R. Botha & M. Everaert (Eds.), *The evolutionary emergence of language: Evidence and Inference* (pp. 42–61).

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(05), 489-558.

Dautriche, I., Mahowald, K., Gibson, E., & Piantadosi, S. T. (2017). Wordform similarity increases with semantic similarity: An analysis of 100 languages. *Cognitive Science*, *41*(8), 2149–2169.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.

Key, M. R., & Comrie, B. (2015). *The Intercontinental Dictionary Series* [Data set]. Retrieved from http://ids.clld.org

Le, Q. V., Jaitly, N., & Hinton, G. E. (2015). A simple way to initialize recurrent networks of rectified linear units. *ArXiv:1504.00941 [Cs]*. Retrieved from http://arxiv.org/abs/1504.00941

List, J. M., Greenhill, S., Rzymski, C., Schweikhard, N., & Forkel, R. (2019). *Concepticon 2.1.0* [Data set]. Retrieved from http://concepticon.clld.org

Monaghan, P., Christiansen, M. H., Farmer, T. A., & Fitneva, S. A. (2010). Measures of phonological typicality: Robust coherence and psychological validity. *The Mental Lexicon*, *5*(3), 281–299.

Vigliocco, G., Vinson, D. P., Druks, J., Barber, H., & Cappa, S. F. (2011). Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. *Neuroscience & Biobehavioral Reviews*, *35*(3), 407–426.

Yujian, L., & Bo, L. (2007). A normalized Levenshtein distance metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(6), 1091–1095.

EVOLUTIONARY PERSPECTIVES ON SPEECH AND INFORMATION RATES

CHRISTOPHE COUPE 1[*1,2], FRANÇOIS PELLEGRINO[2], YOONMI OH[3] and DAN DEDIU[2,4]

[*]Corresponding Author: ccoupe@hku.hk
[1]Department of Linguistics, The University of Hong Kong, Hong Kong, China
[2]Laboratoire Dynamique du Langage, CNRS/University of Lyon, Lyon, France
[3]Department of French Language and Literature, Ajou University, Suwon, South Korea
[4]Collegium de Lyon, University of Lyon, Lyon, France

1. Converging information rates in modern languages

Recent cross-linguistic studies have shown that while languages may differ quite significantly when it comes to speech rate, they are actually much closer in terms of information rate, i.e. the quantity of information they convey on average per second (Coupé et al., 2019). The explanation is a trade-off between speech rate and the average amount of information carried by linguistic units. Syllables have in particular been investigated in different languages and assessed with measures such as conditional entropy (Fenk-Oczlon & Fenk, 2010; Pellegrino, Coupé & Marsico, 2011; Coupé et al., 2019). Mandarin thus has a rather low speech rate, but information-dense syllables, while Spanish has a much faster speech rate, but much lighter syllables. Overall, all languages seem to fall into a narrow range of values for their information rate, centered around 39 bits per second.

The previous result is based on averaged unconstrained speech rates, and points toward universally shared capacities to encode, produce and decode speech. In every linguistic community, however, individual speakers display quite a range of variation (Jacewicz et al., 2010), and can additionally easily increase or decrease their usual speech rate in significant proportions. Unless it becomes extreme, a faster or a slower rate does not impact much the interlocutor's comprehension (Dupoux & Green, 1997; Adank & Janse, 2009). Universal functional and cognitive constraints, although they definitely exist, are therefore relatively weak. An underlying mechanism seems to rest on a coupling between the rhythm of cortical activity and the informational bandwidth of our communication system (Bosker & Ghitza, 2018; Hiafyl et al., 2105).

2. Information rates and speech rates in an evolutionary framework

The previous results can be framed in an evolutionary perspective, and offer insights about how language developed through time, in a way complementary to Villasenor et al. (2012). In this contribution, we distinguish the micro, meso and macro time scales (Wang, 1978) and assess our hypotheses with computational stochastic models of sociolinguistic networks (Nettle, 1999; Gong, Minett, & Wang, 2008), simulating the evolution of communities of speakers under constraints of i) linguistic convergence and ii) 'trade-off' between speech rate and information density.

At the micro scale of years or decades, we investigate how some actuated structural changes may be hindered, despite linguistic convergence, if they take idiolects away from the basin of attraction of information rates, e.g. drastic sound changes in some speakers which strongly increase or decrease the average information density.

At the meso scale of centuries or millennia, although the prehistory of modern languages remains mostly out of reach, language change is visible and occurs in speakers all equipped with modern cognitive capacities. It seems safe to assume information rates similar as today for a long period of time, likely at least since the emergence of our species. For this time period, we explore how minor variations in information density drive the evolution of speech rate, keeping a *fixed* optimal information rate, and how changes can spread from a few speakers to the entire community, if not during one's lifetime, across a few generations. Beyond internally-motivated changes, language contact and borrowing between structurally quite different languages can likely also impact speech rates, with the case of creoles requiring particular attention.

Finally, at a macro time scale, the evolution of the language function itself is characterized by a gradual increase in our ancestors' capacity to convey information, with the development of cognitive abilities and of speech physiological structures. An evolutionary scenario can be sketched, where natural selection is made possible by the aforementioned weak constraints and inter-individual variation, and the selective pressure is the sociocultural development pushing speakers to meet increasing communication needs. We investigate this scenario with a *variable* optimal information rate in our population of artificial speakers, and a general tendency for this rate to increase through time.

References

Adank, P., & Janse, E. (2009). Perceptual learning of time-compressed and natural fast speech. *Journal of the Acoustical Society of America, 126*, 2649-2659.

Bosker, H. R., & Ghitza, O. (2018). Entrained theta oscillations guide perception of subsequent speech: Behavioural evidence from rate normalisation. *Language, Cognition and Neuroscience, 33(8),* 955-967.

Coupé, C., Oh, Y., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: comparable information rates across the human communicative niche. *Science Advances*, *5(9),* eaaw2594.

Dupoux, E., & Green, K. (1997). Perceptual Adjustment to Highly Compressed and Rate Changes Speech: Effects of Talker and Rate Changes. *Journal of Experimental Psychology: Human Perception and Performance*, *23(3),* 914-927.

Fenk-Oczlon, G., & Fenk, A. (2010). Measuring basic tempo across languages and some implications for speech rhythm. In *Proceedings of the 11th Annual Conference of the International Speech Communication Association, Interspeech 2010*, 1537-1540.

Gong, T., Minett, J. W., & Wang, W. S. Y. (2008). Exploring social structure effect on language evolution based on a computational model. *Connection Science*, *20(2–3),* 135–153.

Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A. L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *Elife*, *4,* e06213.

Jacewicz, E. Fox, R. A., O'Neill, C., & Salmons, J. (2009). Articulation rate across dialect, age, and gender. *Language Variation and Change*, *21(2),* 233-256.

Jacewicz, E. Fox, R. A., & Wei, L. (2010). Between-speaker and within-speaker variation in speech tempo of American English. *J. Acoust. Soc. Am.*, *128(2),* 839-580.

Nettle, D. (1999). Using Social Impact Theory to simulate language change. *Lingua*, *108(2–3),* 95–117.

Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language*, *87(3),* 539-558.

Villasenor, J., Han, Y., Wen, D., Gonzales, E., Chen, J., & Wen, J. (2012). The information rate of modern speech and its implications for language evolution. In T. Scott-Phillips, M. Tamariz, E. A. Cartmill, and J. R. Hurford (Eds.), *Evolution of Language, The Proceedings Of The 9th International Conference (Evolang9)* (pp. 376-383). World Scientific.

Wang, W. S-Y. (1978). The Three Scales of Diachrony. In B. B. Kachru (ed.), *Linguistics in the Seventies: Directions and Prospects* (pp.63-76). Department of Linguistics, University of Illinois.

THE CO-EVOLUTION OF COMMUNICATION AND COOPERATION: LEWIS SIGNALLING GAMES IN SOCIAL DILEMMA SCENARIOS

MEGAN CROSS[1*], AMY LUMSDEN[1*], and MATTHEW SPIKE[1†]

[*]Joint first author
[1]Centre for Language Evolution, LEL, PPLS, University of Edinburgh, UK
[†]Corresponding Author: mspike@ed.ac.uk

For communication to emerge and persist in a population, it should be adaptive for both senders and receivers (Skyrms, 2010). However, many social interactions involve conflicts of interest, mis-matches between personal and collective goals, or the temptation to minimize individual risk or free-ride (Sterelny, 2012). Humans are particularly successful at resolving these issues — but as this skill is often attributed to our use of language (Smith, 2010), it appears that language and sociality are evolutionary pre-requisites for each other. This lends weight to theories such as the Interdependence Hypothesis (Tomasello et al., 2012), which argue for the co-evolution of cooperation and communication. However, relatively few accounts of signalling consider more complex social scenarios; modelling work (Santos et al., 2011) has shown that signalling leads to improved group and individual outcomes in repeated *social dilemma* games, where signals serve as 'secret handshakes', i.e. social identifiers, rather than honest signals of intended strategies. Our experimental study investigates which strategies are used by interacting groups of human participants: honest signals, social identifiers, or otherwise.

Methods

112 participants were recruited to play repeated social dilemma games in groups of four, using the online platform oTree (Chen et al., 2016). The two games were i) the *stag hunt*, where individual and common interest coincide, but risk-avoidance strategies are available, and ii) the *prisoner's dilemma*, where defection is individually preferable in any single game, but cooperative strategies are mutually beneficial in the long-term. Groups were allocated to one of four conditions: two conditions without signalling (SH & PD), and two with bidirectional pre-play signalling (SHS & PDS), with 6 graphical signals with no pre-established meaning made available. The four participants within each group were anonymously paired at random for each of 20 rounds of play. Individual and group behaviour was measured as the proportion of cooperative decisions made across rounds; for signals S, decisions D, and identities I, conditional entropy was used to measure the

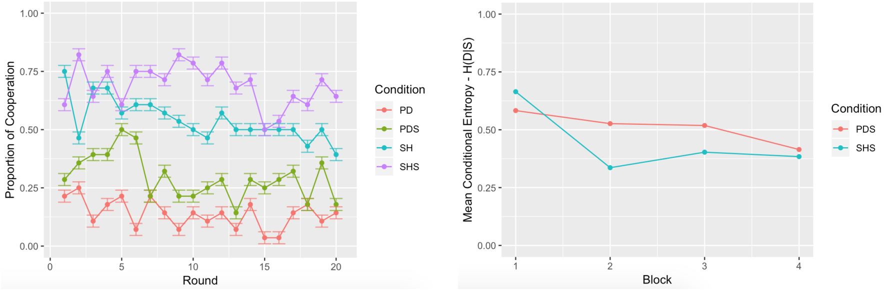


Figure 1. On the left, the mean proportion of cooperation across rounds for the different conditions. On the right, the conditional entropy of decisions given signals $H(D|S)$ across blocked rounds in the SHS (stag hunt with signals) and PDS (prisoner's dilemma with signals) conditions.

informativity over time of $H(S|D)$, $H(D|S)$, $H(S|I)$, and $H(I|S)$.

Results and discussion

1) *Cooperation*: in a linear mixed effects logistic regression with participant nested within group as a random effect, more cooperation was found in stag hunt conditions SH & SHS than in prisoner's dilemma conditions PD & PDS ($\beta = 2.64,\ S.E. = 0.36,\ z = 7.26, p \ll 0.001$; see Fig.1, left). While there was no significant difference between SH and SHS, participants in PDS cooperated more than in PD ($\beta = 0.99,\ S.E. = 0.16,\ z = 6.35, p \ll 0.001$). Finally, cooperation decreased across rounds overall across conditions ($\beta = -0.023,\ S.E. = 0.01,\ z = -2.37, p < 0.05$).

2) *Signal information*: in a linear mixed effects regression including group as a random effect, $H(D|S)$ decreased across blocks ($\chi^2(1) = 4.26, p < 0.05$) by 0.06 ± 0.03 bits per block (see Fig.1, right). The conditional entropy $H(S|D)$ was lower in PDS than SHS ($\chi^2(1) = 14.64, p < 0.001$) by 0.49 ± 0.13 bits. $H(S|I)$ was lower in condition SHS ($\chi^2(1) = 15.92, p < 0.001$ by 0.46 ± 0.11 bits.

The higher levels of cooperation in the signalling conditions supports the general hypothesis that communication promotes cooperation. The slight decrease in cooperation over time across conditions, on the other hand, may be due to the destabilising influence of non-cooperative strategies, to which both scenarios are vulnerable, and potentially even retributive strategies such as punishment and/or spite. However, participants in the signalling conditions did not resort to indiscriminate defection, suggesting the use of more complex coordination strategies. The decrease in $H(D|S)$ combined with the higher levels of cooperation in the signalling conditions implies that, over time, signals served to reduce uncertainty about opponents' strategies. This suggests that, rather than the emergence of social identifiers as predicted by Santos et al (2011), honest signalling plays a role in the development and maintenance of cooperation.

References

Chen, D., Schonger, M., & Wickens, C. (2016). OTree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*(C), 88-97.

Santos, F., Pacheco, J., & Skyrms, B. (2011). Co-evolution of pre-play signaling and cooperation. *Journal of Theoretical Biology*, *274*(1), 30-35.

Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford University Press.

Smith, E. A. (2010). Communication and collective action: language and the evolution of human cooperation. *Evolution and Human Behavior,*, *31*(4), 231-245.

Sterelny, K. (2012). *The evolved apprentice: How evolution made humans unique*. The MIT Press.

Tomasello, M., Melis, A., C. Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation. the interdependence hypothesis. *Current Anthropology*, *53*(6), 673-692.

A GAME-BASED FRAMEWORK FOR STUDYING THE EMERGENCE AND DYNAMICS OF SHARED CONVENTIONS

Christine Cuskley*[1,2] and Simon Kirby[2]

*Corresponding Author: chrstine.cuskley@ncl.ac.uk
[1]Linguistics Centre for Behaviour and Evolution, Newcastle University, Newcastle upon Tyne, UK
[2]Centre for Language Evolution, University of Edinburgh, UK

Gamification in behavioural experiments has become a familiar tool, particularly to study the evolution of language (e.g., Fay, Garrod, Roberts, & Swoboda, 2010; Selten & Warglien, 2007; Irvine & Roberts, 2016). This is mainly approached in two different ways: either by framing the task as a game, introducing game mechanics (e.g., points or rewards), and/or introducing superficial but alluring visual elements which make the task "look" more like a game (Lieberoth, 2015). Recently, Morin et al. (2018) describe the first large-scale attempt to integrate all three of these approaches using *The Colour Game*. Although results are still forthcoming, the game involves a director-matcher style task with colour meanings (represented by swatches) and a pre-specified set of graphical symbols as forms. Much like earlier graphical communication experiments, the explicit goal in the game is to communicate a pre-specified meaning successfully to a partner. However, the scale of this game was much larger, and involved participants being able to choose who they interacted with and how often they played.

Here, we extend efforts in this vein with a multi-player game which revolves around communication with unfamiliar graphical symbols (see Cuskley, 2019 for an example). The game is designed in the style of a browser-based .io game (Castello, 2018) for engaging, voluntary play from the perspective of the participants, and shares broad similarity with agent-based signalling games (Baronchelli, 2016). Players use a pre-specified set of symbols to communicate about colour, but the explicit goal of the game is not communication. Instead, individual players aim to coordinate their own internal colours. Players are embodied 'cells' in a two-dimensional 'petri dish' that can explore freely using simple movements of the mouse. Each cell has a large signal on its body which is visible to other players and can be changed at any time with a simple interface. Players begin as simple cells with only two different coloured 'organelles', and their goal is to become a more complex cell by trading organelles with other players (Figure 1).

This goal is relatively straightforward, but two features have been built into the game which provide a tacit pressure for communication to develop between players. First, the only means by which a player can acquire colours different to

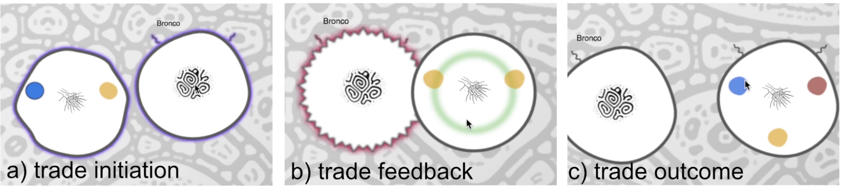


Figure 1. An interaction during game play. All players' central signals are always visible, but players' coloured organelles are only visible to them. Each panel shows a) the initiation of a trade, b) feedback animations for the trade (which resulted in a colour match for the player on the right), and c) the outcome of the trade: the left hand player has levelled up and gained antennae and an additional organelle, while the other player has lost an antenna.

the ones they already have (which is necessary to align their colours and complexify) is to trade organelles with other players. In other words, if they have a red organelle and need a blue one, they must find a player who has a blue one and needs a red one, and propose a trade. Second, a player can only see *their own* organelles, but not the organelles of other players. This means that to make effective colour trades, they need to be able to communicate which colour they are giving away, and seek out other players giving away colours they need. Thus, without the ability to use existing communication channels, players need to develop signalling strategies within the game to coordinate successful trades.

While trade interactions are fundamentally dyadic, the overall game is not: multiple players exist in the same space, and can "overhear" (oversee) interactions between other players. Animations within the game provide information about the outcome of an interaction, and are visible to any players within "eyeshot" of a trade (Fig. 1b). The game records how often players interact, who they interact with, and who else they can see, in addition to the outcomes of interactions for individual players, the signals used, and the meanings (colours) involved. Early pilots of the game with small groups of 4 show two key findings: (i) players report finding game play engaging in and of itself, and their main focus is in-game advancement rather than explicit communciation, and yet, (ii) communication is a pre-requisite for players' success (i.e., players who communicate effectively are more likely to to successfully align their colours, and level up more quickly).

We provide a detailed description of the game concept and mechanics used in these early pilots. We argue that this framework is well-suited not only to studying consensus, but divergence, cooperation and competition, naturalistic social network structure, and issues surrounding form and meaning space constraints. In short, we demonstrate specific potential for studying the emergence of language-like conventions, but also provide an open-ended 'petri dish' for experiments in cultural evolution and collective behaviour generally.

Acknowledgements

This research was supported by British Academy Postdoctoral Fellowship grant no. pf150065.

References

Baronchelli, A. (2016). A gentle introduction to the minimal naming game. *Belgian Journal of Linguistics*, *30*(1), 171–192.

Castello, J. (2018). The rise and rise of io games. In *Rock paper shotgun.* `https://www.rockpapershotgun.com/2018/02/22/the-rise-and-rise-of-io-games/`, accessed 2019-09-13.

Cuskley, C. (2019). Alien forms for alien language: investigating novel form spaces in cultural evolution. *Palgrave Communications*, *5*(1), 1–15.

Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive science*, *34*(3), 351–386.

Irvine, E., & Roberts, S. (2016). Deictic tools can limit the emergence of referential symbol systems. In S. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 11th international conference (evolangx11).* Online at `http://evolang.org/neworleans/papers/99.html`.

Lieberoth, A. (2015). Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture*, *10*(3), 229–248.

Morin, O., Winters, J., Müller, T. F., Morisseau, T., Etter, C., & Greenhill, S. J. (2018). What smartphone apps may contribute to language evolution research. *Journal of Language Evolution*, *3*(2), 91–93.

Selten, R., & Warglien, M. (2007). The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences*, *104*(18), 7361–7366.

THE GENERALITY OF FORM-TO-MEANING ICONICITY

CHARLES P. DAVIS*[1], HANNAH M. MORROW[1], MARGHERITA DE LUCA[2], and GARY LUPYAN[3]

*Corresponding Author: charles.davis@uconn.edu
[1]Department of Psychological Sciences, University of Connecticut, Storrs, USA
[2]Department of Document Studies, Linguistics, Philology and Geography, Sapienza University of Rome, Rome, Italy
[3]Department of Psychology, University of Wisconsin-Madison, Madison, USA

Do certain words fit some meanings better than others? If so, to what extent do people make use of these relationships? Given a choice of whether a "nonsense" word refers to one or another object, people show predictable choices (e.g., Sapir, 1929; see Dingemanse et al., 2015 for review). Understanding the generality and origins of iconicity – a resemblance between a word's form and its meaning – is critical for research in language evolution as it lends insight into how initial form-to-meaning links may be established prior to emergence of large-scale conventional vocabularies.

To test whether form–meaning resemblances affect behavior in a relatively open-ended task, in a recent study (Davis et al., 2019), we first asked people to match written English-like nonsense words (from Westbury et al., 2018) to properties, e.g. is a *horgous* large, round, etc. (Fig. 1). A second group drew creatures in response to the same nonce words. A third group was asked to indicate whether the drawn creatures were large, round, etc. A final group was shown the creatures and asked to match them to nonsense words. Remarkably, the form of the nonsense words permeated the creature drawings: people judged the drawn creatures as having the same properties connoted by the original nonce words and were able to match the drawn creatures back to the name used to elicit them with higher-than-chance accuracies (e.g. matching the *horgous* drawings back to the word "horgous").

What explains these associations? One possibility is that they are mediated by idiosyncratic similarity to real English words, e.g., a "horgous" may be *large* due to form overlap with "humongous." Alternatively, the form–meaning associations may reflect more general sources of information, e.g., experience associating perceptual inputs across modalities (Lupyan & Casasanto, 2015).

To distinguish between these possibilities, we recruited native Spanish and Italian speakers to match the nonce words to translated properties. Fig. 1 shows how similar the results were across languages (cross-language correlations: $r = .75–.91$). Some discrepancies (e.g., *cougzer* as feminine vs. masculine) likely stem from overlap with real-word neighbors, but these do not explain the overwhelming similarity, as the orthographic neighborhoods of the nonce words in each language differ considerably.

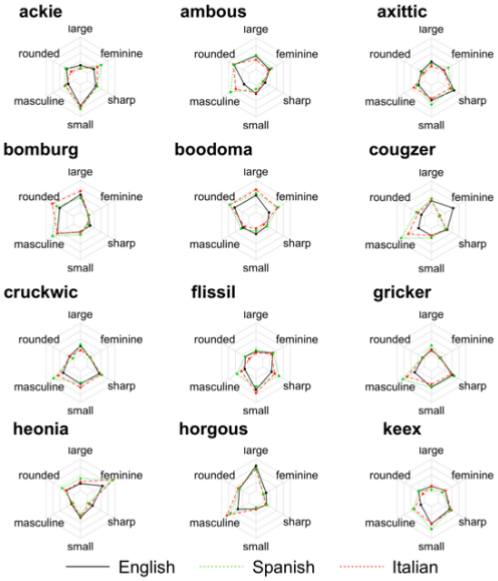


Fig. 1. Nonce-word property ratings for three languages.

We next examined whether the form–meaning associations are present in the distributed structure of each language. We trained word embedding models using the fast-text algorithm (Bojanowski et al., 2016) trained on English, Spanish, and Italian Wikipedia as well as parallel-translations of the OpenSubtitles corpus. We used the nonce words as input to the model and examined the proximity of resulting semantic representations to each property (e.g., *large*). The models were correlated with human ratings ($z > 10$). Strikingly, predictions were stronger for *between*-language pairs (e.g., English responses were more poorly predicted by English embeddings than by Spanish and Italian embeddings). This is unexpected if iconic associations derive from nearby real words, as such proximity-driven responding would increase within-language matches.

Nonarbitrary form–meaning associations appear to be surprisingly potent, influencing even open-ended drawing tasks. When matching nonce words to properties, English, Italian, and Spanish speakers show strikingly similar behavior (though replication to more diverse languages is clearly needed, as the languages here were chosen for convenience). Lastly, behavior was predicted by large-scale form–meaning associations in language as learned by a simple neural network.

References

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146.

Davis, C. P., Morrow, H. M., & Lupyan, G. (2019). What does a horgous look like? Nonsense words elicit meaningful drawings. *Cognitive Science*, *43*(10), e12791.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.

Lupyan, G., & Casasanto, D. (2015). Meaningless words promote meaningful categorization. *Language and Cognition*, *7*(2), 167–193.

Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology*, *12*, 225–239.

Westbury, C., Hollis, G., Sidhu, D. M., & Pexman, P. M. (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language*, *99*, 122–150.

ANALYSIS OF MUTATION AND FIXATION FOR LANGUAGE

BART DE BOER*[1], BILL THOMPSON[2], ANDREA RAVIGNANI[1,3],
and CEDRIC BOECKX[4]

*Corresponding Author: bart@ai.vub.ac.be
[1]AI-lab, Vrije Universiteit Brussel, Brussels, Belgium
[2]Social Science Matrix, UC Berkeley, Berkeley, USA
[3]Seal Centre, Pieterburen, the Netherlands
[4]ICREA and Universitat de Barcelona, Barcelona, Spain

Did the biological basis for language emerge instantaneously from a single mutation (Berwick & Chomsky, 2016) or gradually through many (Hurford, 2014)? Advances in what we know about how humans have evolved make it possible to formally analyze these alternative evolutionary scenarios. Evolution operates through transmission, variation and selection. Theoretical biology has developed a set of (mathematical) tools with which these processes can be studied (see Blythe & McKane, 2007 for a review relevant to language evolution). These can be used to assign likelihoods to different evolutionary scenarios, and also to calculate estimates of how much time would be needed for each scenario.

The two relevant techniques are: 1) a way to calculate probabilities of mutations and 2) a way to calculate the probability that a mutation spreads in the population (in biological parlance: goes to fixation), as well as the time it would take to spread. Multiplying these probabilities gives the probability of a mutation occurring and subsequently going to fixation. The first is provided by extreme value theory applied to the probability of beneficial mutations, developed by Orr (2003). From this it follows that the probability of a beneficial mutation is approximately $p_{mut}(\varphi)=\alpha\cdot e^{-\alpha\cdot\varphi}$, where φ is the fitness effect of the beneficial mutation (i.e. how much more offspring an organism with the mutation has on average than an organism without it) and α a constant, with a value that is at least 30 (de Boer, Thompson, Ravignani, & Boeckx, to appear).

The probability of fixation and the time to fixation have been worked out by among others Kimura (Kimura, 1957; Kimura & Ohta, 1969). The fixation probability is given by: $p_{fix}(\varphi)=(1-e^{\beta})/(1-e^{\beta \cdot N})$, where $\beta = -4\varphi/(2+\varphi)$ and N is the effective population size. The expression for the fixation time can be determined as well, but is too complicated to include here.

Combining these it is found that in order to achieve a large increase of functionality (approximated by the fitness) the most likely scenario consists of a medium number of mutations with a medium fitness effect (i.e around 0.1). Larger mutations are exceedingly rare, while smaller mutations tend not to reach fixation in the population. This is in line with what has been proposed by biologists (Eyre-Walker & Keightley, 2007). The precise relation between the istribution of beneficial mutations, α, the fitness improvement to reach, *I* and the needed number of mutations *m*, is given in figure 1. Estimating the duration of these scenarios is still a challenge, but even small mutations reach fixation relatively quickly – if it does not disappear from the population – and the time it takes for a mutation to reach fixation depends more on population size than on the strength of the fitness effect. Multiple mutations can evolve in parallel, so the time it takes is about equal to a small multiple of the time it takes for the slowest mutation to reach fixation.

Among other things, these results show that the minimalist account of language is not supported by evolutionary theory. Minimalism was proposed because it was assumed language emerged suddenly and recently. However, our analyses show that even a multi-mutation scenario does not need to take a long time. In other words, complex traits can evolve rapidly if population sizes are not too large, and if the selection pressures remain constant.

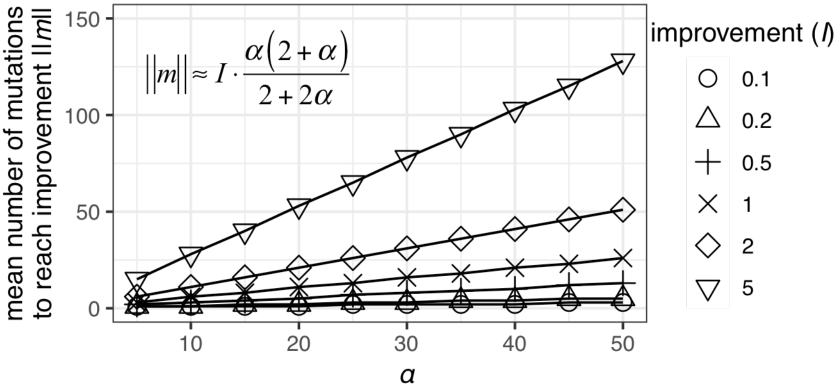


Figure 1. Most likely number of mutations $||m||$ needed to reach a fitness improvement I for different values of the mutation parameter α. Realistic values of α are higher than 30.

References

Berwick, R. C., & Chomsky, N. (2016). *Why only us: Language and evolution*. MIT press.

Blythe, R. A., & McKane, A. J. (2007). Stochastic models of evolution in genetics, ecology and linguistics. *Journal of Statistical Mechanics: Theory and Experiment*, *2007*(07), P07018.

de Boer, B., Thompson, B., Ravignani, A., & Boeckx, C. (to appear). Evolutionary Dynamics Do Not Motivate a Single-Mutant Theory of Human Language. *Scientific Reports*.

Eyre-Walker, A., & Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, *8*(8), 610.

Hurford, J. R. (2014). *The Origins of Language: A Slim Guide*. Oxford: Oxford University Press.

Kimura, M. (1957). Some problems of stochastic processes in genetics. *The Annals of Mathematical Statistics*, *28*(4), 882–901.

Kimura, M., & Ohta, T. (1969). The Average Number of Generations until Fixation of a Mutant Gene in a Finite Population. *Genetics*, *61*(3), 763–771. Retrieved from PMC. (PMC1212239)

Orr, H. A. (2003). The distribution of fitness effects among beneficial mutations. *Genetics*, *163*(4), 1519–1526.

RHYTHMIC PERCEPTION, PRODUCTION AND INTERACTIVITY IN HARBOUR AND GREY SEALS

MARIANNE DE HEER KLOOTS*[1,2], DARYLL CARLSON[2,3], MAXIME GARCIA[4],
SONJA A. KOTZ[5], ALICE LOWRY[2], LAETITIA POLI-NARDI[2], KOEN DE REUS[2,6],
ANA RUBIO-GARCÍA[2], MARLENE SROKA[2], MILA VAROLA[2],
ANDREA RAVIGNANI[1,2,7,8]

*Corresponding Author: mariannedhk@gmail.com
[1]Institute for Advanced Study, University of Amsterdam, Amsterdam, The Netherlands
[2]Research Department, Sealcentre Pieterburen, Pieterburen, The Netherlands
[3]Department of Earth Systems Science, Stanford University, Stanford, United States
[4]Department of Evolutionary Biology and Environmental Studies,
University of Zürich, Zürich, Switzerland
[5]Neuropsychology and Psychopharmacology, Maastricht University,
Maastricht, The Netherlands
[6]Department Life Sciences, Erasmus University Rotterdam, Rotterdam, The Netherlands
[7]Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics,
Nijmegen, The Netherlands
[8]Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium

Comparative studies on rhythmic animal behaviour can help understand the evolutionary origins of rhythm cognition underlying many human behaviours, including speech (Kotz, Ravignani & Fitch, 2018) and conversational interaction (Pika, Wilkinson, Kendrick & Vernes, 2018). Pinnipeds are a particularly promising clade for comparative investigations in the vocal domain (Ravignani et al., 2016), showing remarkable vocal flexibility (Ralls et al., 1985; Reichmuth & Casey, 2014) as well as rhythmic capacities (Cook et al., 2013; Mathevon et al., 2017). The variability in these traits across species guides hypotheses from evolutionary neuroscience, that postulate mechanistic connections between vocal learning and rhythm (Kuypers, 1958a, 1958b; Jürgens, Kirzinger, & von Cramon, 1982; Patel & Iversen, 2014). In seal pups, who are born and weaned in large breeding colonies on land, vocal rhythmicity can be functionally explained by their socio-ecology: a pup calling for its mother's attention may avoid acoustic masking by vocalizing in turns with nearby conspecifics (Ravignani, 2018). Here we disentangle different aspects of seal pup rhythm cognition to better describe individual timing patterns in turn-taking choruses. We present data from four studies on harbour seals (*Phoca vitulina*) and grey seals (*Halichoerus grypus*), testing rhythm perception, production, and interactivity.

In the first study, we tested 20 wild-born harbour seals regarding their capacity to distinguish rhythmic properties in a listening task. The experiment used a set-up inspired by the head-turn preference procedure from infant development studies (Nelson et al., 1995). We measured the number of looks and look duration of individual seals towards the playback stimulus, comparing experimental manipulations for 4 properties: rhythmic regularity (random vs. isochronous), tempo (slow vs. fast), duration (long vs. short calls) and sex (calls produced by a female vs. a male pup). Regularity, tempo, and duration significantly affected the behavioural measures. These results show the importance of rhythmicity as a call property affecting harbour seals' behavioural responses, and support their fine-grained perceptual discrimination capacities in the temporal domain.

The second study (Ravignani et al., 2018) examined rhythmicity in spontaneously produced calls by 3 individual wild-born harbour seal pups, exploring the ontogeny of temporal call organization over the course of several weeks. Analyses on call duration, inter-onset interval, and inter-peak interval showed that call rhythms evolve, with some properties becoming more categorically structured over time. These results demonstrate harbour seals' productive rhythmic flexibility in development and suggest a vocal production system particularly suitable for interactive settings.

Next, we tested the rhythmic interactivity of harbour seal pup vocalizations in a playback experiment eliciting vocal responses. Manipulating regularity, tempo, and identity of the playback calls, we measured the absolute time and relative phase of individual pups' responses. We contrasted the experimental data with results from agent-based simulations to compare several alternative call timing strategies. The results show that seal pups interactively time their vocal responses to conspecific calls, avoiding overlap by taking turns.

Our final study investigates rhythmic interactions in spontaneous settings, using similar analyses in a different species. We analyzed interactive call timing in multitrack recordings of grey seal pups housed in groups. By studying groups of multiple individuals, we applied the analysis techniques from the previous studies in a more ecological setting, and explored how individual call timing strategies contribute to a group-level chorus.

We conclude that seals both perceive and produce vocal rhythmicity, without explicit training. Our results suggest that seal pups interactively time their calls, such that they vocalize in turns, maximizing individual conspicuousness by minimizing acoustic masking. Both harbour and grey seal pup vocalizations are known to have individually distinctive spectral properties, which are used in mother-offspring recognition (Perry & Renouf, 1988; McCulloch & Boness, 1999). We argue that socio-ecological pressures for individual conspicuousness and distinguishability shape both the spectral and temporal dimensions of seal pup vocal communication. Pinnipeds provide a useful model for convergently evolved rhythmic traits, and our results might inform future cross-species work on timing in communicative behaviours — including human speech.

Acknowledgements

We thank the staff and volunteers at Sealcentre Pieterburen for their help during data collection.

References

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology, 127*(4), 412–427. https://doi.org/10.1037/a0032345

Jürgens, U., Kirzinger, A., & von Cramon, D. (1982). The Effects of Deep-Reaching Lesions in the Cortical Face Area on Phonation a Combined Case Report and Experimental Monkey Study. *Cortex, 18*(1), 125–139. https://doi.org/10.1016/S0010-9452(82)80024-5

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The Evolution of Rhythm Processing. *Trends in Cognitive Sciences, 22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002

Kuypers, H. G. J. M. (1958a). Corticobulbar connections to the pons and lower brainstem in man: An anatomical study. *Brain, 81*(3), 364–388. https://doi.org/10.1093/brain/81.3.364

Kuypers, H. G. J. M. (1958b). Some projections from the peri-central cortex to the pons and lower brain stem in monkey and chimpanzee. *Journal of Comparative Neurology, 110*(2), 221–255. https://doi.org/10.1002/cne.901100205

McCulloch, S., & Boness, D. J. (2000). Mother–pup vocal recognition in the grey seal (*Halichoerus grypus*) of Sable Island, Nova Scotia, Canada. *Journal of Zoology*, 251(4), 449–455.

Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern Elephant Seals Memorize the Rhythm and Timbre of Their Rivals' Voices. *Current Biology, 27*(15), 2352-2356.e2. https://doi.org/10.1016/j.cub.2017.06.035

Nelson, D. G. K., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant behavior and development, 18*(1), 111-116.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*. https://doi.org/10.3389/fnsys.2014.00057

Perry, E. A., & Renouf, D. (1988). Further studies of the role of harbour seal (Phoca vitulina) pup vocalizations in preventing separation of mother–pup pairs. *Canadian Journal of Zoology, 66*(4), 934–938. https://doi.org/10.1139/z88-138

Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: Bridging the gap between human and animal communication. *Proceedings of the Royal Society B: Biological Sciences, 285*(1880), 20180598. https://doi.org/10.1098/rspb.2018.0598

Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, Phoca vitulina. *Canadian Journal of Zoology, 63*(5), 1050–1056. https://doi.org/10.1139/z85-157

Ravignani, A., Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., … de Boer, B. (2016). What Pinnipeds Have to Say about Human Speech, Music, and the Evolution of Rhythm. *Frontiers in Neuroscience, 10.* https://doi.org/10.3389/fnins.2016.00274

Ravignani, A., Kello, C. T., de Reus, K., Kotz, S. A., Dalla Bella, S., Méndez-Aróstegui, M., … de Boer, B. (2018). Ontogeny of vocal rhythms in harbor seal pups: An exploratory study. *Current Zoology*, zoy055. https://doi.org/10.1093/cz/zoy055

Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current Opinion in Neurobiology, 28*, 66–71. https://doi.org/10.1016/j.conb.2014.06.011

RELATIONSHIPS BETWEEN VOCAL ONTOGENY AND VOCAL TRACT ANATOMY IN HARBOUR SEALS (*PHOCA VITULINA)*

KOEN DE REUS*[1,2], DARYLL CARLSON[2,3], YANNICK JADOUL[4], ALICE LOWRY[2], STEPHANIE GROSS[5], MAXIME GARCIA[6], ANNA SALAZAR CASALS[2], ANA RUBIO-GARCIA[2], CAROLINE ELISABETH HAAS[7], BART DE BOER[4], ANDREA RAVIGNANI[2,4,8]

*Corresponding Author: kdereus95@gmail.com
[1]Department Life Sciences, Erasmus University College, Erasmus University Rotterdam, Rotterdam, The Netherlands
[2]Research Department, Sealcentre Pieterburen, Pieterburen, The Netherlands
[3]Department of Earth Systems Science, Stanford University, Stanford, USA
[4]Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium
[5]Institute for Terrestrial and Aquatic Wildlife Research (ITAW), University of Veterinary Medicine Hannover, Foundation, Büsum, Germany
[6]Animal Behaviour, Department of Evolutionary Biology and Environmental Science, University of Zurich, Zurich, Switzerland
[7]Cognitive Ethology Laboratory, German Primate Centre, Göttingen, Germany
[8]Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Understanding the origins and evolution of human speech benefits from a multidisciplinary and comparative approach. Research on animal models has already provided some valuable insight into the biological underpinnings of vocal communication. One important focus in the field of animal communication is sound production. The current literature on this topic hosts a great number of studies on avian species and our closer relatives, non-human primates. However, many pinniped species have been reported to have wide vocal repertoires, often producing call types in specific behavioral contexts (Ralls et al., 1985; Mathevon et al., 2017; Charrier et al., 2009). In fact, the vocal abilities of pinnipeds are better than was historically believed (Ravignani et al., 2016). Moreover, pinnipeds are phylogenetically closer to humans than the well-studied birds (O'Leary et al, 2013) and share some anatomical similarities to the human vocal apparatus (Fitch, 2000). Here, we: (i) report on longitudinal data on vocal ontogeny in harbor seal pups, (ii) complement the bioacoustic findings with results from a large anatomical data set of larynges, and (iii) critically compare our findings with available literature on harbor seal sound production. Taken together, they suggest that phocids are good candidates for animal models in future research on the evolution of human speech. In harbor seals (*Phoca vitulina*), young pups produce mother attraction calls which play an important role in parent-offspring recognition. Female reproductive success relies almost entirely on the accurate identification of offspring within the colony after foraging trips (Insley et al., 2003). Studies conducted on harbor seal mother attraction calls found that these calls are individually distinctive (Renouf, 1984; Perry & Renouf, 1988) and that mothers can recognize the calls of their offspring shortly after birth (Renouf, 1985). It is still uncertain which call parameters are used by the mother to recognize the pup and whether she regularly needs to update her acoustic template of the pup vocalisations. In fact, previous studies on the ontogeny of mother attraction calls have shown that this call type changes in acoustic (Khan et al., 2006; Sauvé et al., 2015) and temporal structure (Ravignani et al., 2018), with potential for vocal production learning. Using longitudinal call data, we found that many call features covary with age and sex, but not with body length. Our findings also further provide evidence for the importance of the fundamental frequency and its modulation as individualised call parameters that could play an important role in successful parent-offspring recognition. As pups grow, the development of the anatomical structures making up the vocal tract will impose constraints on the acoustic signals produced. The acoustic allometry framework suggests that vocal tract length is constrained by body size, but vocal fold length is not (confirmed in seals using anatomical data; Ravignani et al. 2017). In light of the source-filter theory, acoustic parameters shaped by the filter such as formants would then be better at conveying body size information than the fundamental frequency, a source parameter (Fant, 1960). Our anatomical measurements of harbor seal vocal tract and larynx indicate that neither vocal tract nor vocal fold length correlate with body size. Body size may be a fundamental driver of acoustic signals (Martin et al., 2017), but future allometry studies will shed more light on which call features accurately encode size information and how these interact with vocal development, plasticity and learning.

Acknowledgements

We thank all staff and volunteers from Sealcentre Pieterburen for their help in the data collection, hospitality and continued support in research.

References

Charrier, I., Pitcher, B. J., & Harcourt, R. G. (2009). Vocal recognition of mothers by Australian sea lion pups: individual signature and environmental constraints. *Animal Behaviour*, *78*(5), 1127-1134.

Fant G. (1960). Acoustic Theory of Speech Production. The Hague, The Netherlands: Mouton & Co. Publishers.

Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, *4*, 258-267.

Insley, S. J., Phillips, A. V., & Charrier, I. (2003). A review of social recognition in pinnipeds. *Aquatic Mammals*, *29*(2), 181-201.

Khan, C. B., Markowitz, H., & McCowan, B. (2006). Vocal development in captive harbor seal pups, Phoca vitulina richardii: Age, sex, and individual differences. *The Journal of the Acoustical Society of America*, *120*(3), 1684-1694.

Martin, K., Tucker, M. A., & Rogers, T. L. (2017). Does size matter? Examining the drivers of mammalian vocalizations. *Evolution*, *71*(2), 249-260.

Mathevon, N., Casey, C., Reichmuth, C., & Charrier, I. (2017). Northern elephant seals memorize the rhythm and timbre of their rivals' voices. *Current Biology*, *27*(15), 2352-2356.

O'Leary, M.A., Bloch, J.I., Flynn, J.J., Gaudin, T.J., Giallombardo, A., Giannini, N.P., Goldberg, S.L., Kraatz, B.P., Luo, Z.X., Meng, J. and Ni, X. (2013). The placental mammal ancestor and the post–K-Pg radiation of placentals. *Science*, *339*(6120), 662-667.

Perry, E. A., & Renouf, D. (1988). Further studies of the role of harbour seal (Phoca vitulina) pup vocalizations in preventing separation of mother–pup pairs. *Canadian Journal of Zoology*, *66*(4), 934-938.

Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, Phoca vitulina. *Canadian Journal of Zoology*, *63*(5), 1050-1056.

Ravignani, A., Fitch, W., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., Scharff, C., Stoeger, A. & de Boer, B. (2016). What pinnipeds have to say about human speech, music, and the evolution of rhythm. *Frontiers in Neuroscience*, *10*, 274.

Ravignani, A., Gross, S., Garcia, M., Rubio-Garcia, A., & de Boer, B. (2017). How small could a pup sound? The physical bases of signaling body size in harbor seals. *Current Zoology*, *63*(4), 457-465.

Ravignani, A., Kello, C. T., de Reus, K., Kotz, S. A., Dalla Bella, S., Méndez-Aróstegui, M., Rapado, B., Rubio-Garcia, A. & de Boer, B. (2018). Ontogeny of vocal rhythms in harbor seal pups: an exploratory study. *Current Zoology*, *65*(1), 107-120

Renouf, D. (1984). The vocalization of the harbour seal pup (Phoca vitulina) and its role in the maintenance of contact with the mother. *Journal of Zoology*, *202*(4), 583-590.

Renouf, D. (1985). A demonstration of the ability of the harbour seal PhocaVitulina (L.) to discriminate among pup vocalizations. *Journal of Experimental Marine Biology and Ecology*, *87*(1), 41-46.

Sauvé, C. C., Beauplet, G., Hammill, M. O., & Charrier, I. (2015). Acoustic analysis of airborne, underwater, and amphibious mother attraction calls by wild harbor seal pups (Phoca vitulina). *Journal of Mammalogy*, *96*(3), 591-602.

VOCAL TRACT ANATOMY AS A TIME MACHINE: WHAT CAN WE INFER ABOUT THE “DEEP” LINGUISTIC PAST?

DAN DEDIU*[1] and SCOTT R. MOISIK[2]

*Corresponding Author: dan.dediu@univ-lyon2.fr
[1]Laboratoire Dynamique Du Langage (DDL), Université Lumière Lyon 2, Lyon, France
[2]Linguistics and Multilingual Studies, Nanyang Technological University,
Singapore, Singapore

It is usually thought that languages change too fast for them to retain any information about events and processes in the “deep” past, with those beyond a “time horizon” of about 10,000 years or so before the present being effectively lost (Renfrew et al., 2000). While this might well be the case for “classical” historical linguistics (Campbell, 2004; Lass, 1997), new (and sometimes quite controversial) methods and data seem to suggest otherwise. As such, modern Bayesian phylogenetic methods borrowed from evolutionary biology applied to cognacy judgments on basic vocabulary lists (Atkinson & Gray, 2005) have allowed detailed reconstructions of a few large language families (such as *Indo-European* and *Austronesian*) which may go back 7-8,000 years before present (Bouckaert et al., 2012; Gray & Atkinson, 2003; Honkola et al., 2013). More controversially, similar methods have been recently used to suggest that some Eurasian language families might be related through a shared ancestor some 15,000 years ago (Pagel et al., 2013), a suggestions also supported by a different approach using phylogenetic inferences on the alignment of actual transcriptions from the ASJP database (Jäger, 2015). A very indirect support may be offered by between-families patterning of the stability of structural features of language, suggesting not only that some Eurasian families were connected in the “deep” past, but also that the language families of America and north-eastern Eurasia might have been linked, presumably 15,000 years ago or so (Dediu & Levinson, 2012). However, while these new findings may be able push back the “time horizon”, it still concerns but a sliver of the at least 500,000 years or so of modern language and speech (Dediu & Levinson, 2013, 2018).

In this talk, we propose that recent work linking features of vocal tract anatomy with phonetics and phonology might allow inferences about some features of languages spoken in the “deep” past. First, we will review evidence that some of the hard structures of the vocal tract can be recovered from the osteological and fossil records in fine enough detail to allow such inferences to be made, focusing on the *hyoid bone* (Martínez et al., 2008), the *lower jaw* (Bosman et al., 2017) and the *hard palate* (Baetsen, 2016). Second, we will use recently published and ongoing work to show that metric variation in particular features of the vocal tract (using both “classic” measurements, such as *distances and angles*, and *geometric morphometric* approaches separating shape from size; Zelditch et al., 2012), showing inter-individual and inter-group variation, do influence speech. We focus on the effects of the *alveolar ridge* on click consonants (Moisik & Dediu, 2017), of *bite* on labiodentals (Blasi et al., 2019), of the *hard palate* on [ɹ] (Dediu & Moisik, 2019), and of *larynx position* and *hard palate midsagittal shape* on vowels (Dediu et al., 2019; Janssen et al., 2019). Finally, work in the tradition of the iterated learning model (ILM) suggests that such weak biases may be amplified by the repeated use and learning of language (Dediu et al., 2019; Kirby et al., 2007), and that such anatomical variation might indeed affect sound change (Dediu & Moisik, 2019) either directly (through its immediate articulatory and acoustic effects) or indirectly (by changing the probabilities of sound change elsewhere through covert coarticulatory influences).

Taken together, these suggest that, by looking at the osteological and fossil record, we may be able to make probabilistic inferences about the languages of the past. Thus, we can pretty safely infer that obstruents closer to the glottis were less likely to be voiced (Everett, 2018) as far back as speech existed (and that this tendency might have been even stronger for higher larynx positions), and that labiodentals were much less frequent before the spread of agriculture (Blasi et al., 2019). Other inferences require more data that is currently available, such as the hard palate shape of populations in Africa before the Bantu expansion (were clicks more widespread there at that time?) and elsewhere (were clicks present? and, if any, what variations of [ɹ] would be most prevalent?). More importantly however, is that such studies allow quantitative, empirical inferences about the past.

Acknowledgements

This work was supported by funding from the Dutch Research Council (NWO 276-70-022), the European Institutes for Advanced Study (EURIAS 2017-2018) and the French National Research Agency (ANR)/IDEXLyon (16-IDEX-0005).

References

Atkinson, Q. D., & Gray, R. D. (2005). Curious parallels, curious connections—Phylogenetic Thinking in Biology and Historical Linguistics. *Systematic Biology*, *54*(4), 513–526.

Baetsen, W. A. (2016). *Geometric morphometric analysis of the "skeletal vocal tract": A first step in involving osteoarchaeology in the search for a potential "genetic bias" for language, using two Dutch historical skeletal populations* (MSc Thesis). University of Leiden, Leiden, The Netherlands.

Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science*, *363*(6432), eaav3218. https://doi.org/10.1126/science.aav3218

Bosman, A. M., Moisik, S. R., Dediu, D., & Waters-Rist, A. (2017). Talking heads: Morphological variation in the human mandible over the last 500 years in the Netherlands. *HOMO - Journal of Comparative Human Biology*, *68*(5), 329–342. https://doi.org/10.1016/j.jchb.2017.08.002

Bouckaert, R., Lemey, P., Dunn, M., Greenhill, S. J., Alekseyenko, A. V., Drummond, A. J., … Atkinson, Q. D. (2012). Mapping the Origins and Expansion of the Indo-European Language Family. *Science*, *337*(6097), 957–960. https://doi.org/10.1126/science.1219669

Campbell, L. (2004). *Historical linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Dediu, D., Janssen, R., & Moisik, S. R. (2019). Weak biases emerging from vocal tract anatomy shape the repeated transmission of vowels. *Nature Human Behaviour*, *3*, 1107–1115. https://doi.org/10.1038/s41562-019-0663-x

Dediu, D., & Levinson, S. C. (2012). Abstract Profiles of Structural Stability Point to Universal Tendencies, Family-Specific Factors, and Ancient Connections between Languages. *PLoS ONE*, *7*(9), e45198. https://doi.org/10.1371/journal.pone.0045198

Dediu, D., & Levinson, S. C. (2013). On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Language Sciences*, *4:*, 397. https://doi.org/10.3389/fpsyg.2013.00397

Dediu, D., & Levinson, S. C. (2018). Neanderthal language revisited: Not only us. *Current Opinion in Behavioral Sciences*, *21*, 49–55. https://doi.org/10.1016/j.cobeha.2018.01.001

Dediu, D., & Moisik, S. R. (2019). Pushes and pulls from below: Anatomical variation, articulation and sound change. *Glossa: A Journal of General Linguistics*, *4*(1), 7. https://doi.org/10.5334/gjgl.646

Everett, C. (2018). The global dispreference for posterior voiced obstruents: A quantitative assessment of word-list data. *Language*, *94*(4), e311–e323. https://doi.org/10.1353/lan.2018.0069

Gray, R. D., & Atkinson, Q. D. (2003). Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*, *426*, 435–439.

Honkola, T., Vesakoski, O., Korhonen, K., Lehtinen, J., Syrjänen, K., & Wahlberg, N. (2013). Cultural and climatic changes shape the evolutionary history of the Uralic languages. *Journal of Evolutionary Biology*, *26*(6), 1244–1253. https://doi.org/10.1111/jeb.12107

Jäger, G. (2015). Support for linguistic macrofamilies from weighted sequence alignment. *Proceedings of the National Academy of Sciences*, *112*(41), 12752–12757. https://doi.org/10.1073/pnas.1500331112

Janssen, R., Moisik, S. R., & Dediu, D. (2019). The effects of larynx height on vowel production are mitigated by the active control of articulators. *Journal of Phonetics*, *74*, 1–17. https://doi.org/10.1016/j.wocn.2019.02.002

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proc Natl Acad Sci U S A*, *104*(12), 5241–5245. https://doi.org/10.1073/pnas.0608222104

Lass, R. (1997). *Historical linguistics and language change*. Cambridge : Cambridge University Press.

Martínez, I., Arsuaga, J. L., Quam, R., Carretero, J. M., Gracia, A., & Rodríguez, L. (2008). Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). *Journal of Human Evolution*, *54*(1), 118–124. https://doi.org/10.1016/j.jhevol.2007.07.006

Moisik, S. R., & Dediu, D. (2017). Anatomical biasing and clicks: Evidence from biomechanical modeling. *Journal of Language Evolution*, *2*(1), 37–51. https://doi.org/doi:10.1093/jole/lzx004

Pagel, M., Atkinson, Q. D., Calude, A. S., & Meade, A. (2013). Ultraconserved words point to deep language ancestry across Eurasia. *Proceedings of the National Academy of Sciences*, *110*(21), 8471–8476. https://doi.org/10.1073/pnas.1218726110

Renfrew, C., McMahon, A., & Trask, L. (Eds.). (2000). *Time depth in historical linguistics*. McDonald Institute, UK.

Zelditch, M. L., Swiderski, D. L., & Sheets, H. D. (2012). *Geometric Morphometrics for Biologists: A Primer*. Amsterdam: Academic Press.

A COMPARATIVE APPROACH TO THE LARYNGEAL SOUND SOURCE OF PRIMATES: IMPLICATIONS FOR THE EVOLUTION OF SPEECH

Didier Demolin
Laboratoire de phonétique et phonologie, CNRS-UMR 7018, Sorbonne nouvelle
didier.demolin@sorbonne-nouvelle.fr

Human beings have the capacity of rapid changes in the shape of their vocal tract during speech production. This allows the production of formants providing the acoustic basis of vowels. Much research has been done to identify the capacity of non-primates to produce vowels (e.g. Fitch et al. 2016). However it would be erroneous to assume that the sound production of primates is limited to resonances in the vocal tract. There are many researches into the call of monkeys and apes showing that various species produce complex patterns of vocalizations for communication. The comparative anatomy and physiology of the monkeys and apes larynx reveal many important features on the sound source. The monkeys and apes larynges differ one from another and those of humans in size, intranarial position and also by the presence of air sacs. There is little doubt that most monkeys and apes produce laryngeal sound source for specific purposes. Grunts hoot barks, pant-hoot, loud calls and chuckles are well known. Current views suggest that the non-human primate source generates an unstable glottal source. Some observations on the shape and structure of the vocal folds by Schön Ybara (1995) on platyrrhines, Starck and Schneider (1960) on Chimpanzees and Hirano (1991) reveal some characteristic features and the sound-producing capabilities of the non-human primate vocal folds. According to Schön Ybarra (1995), the non-human primate larynx appears to have more phonatory range, but less phonatory precision that that of humans. The comparison the acoustic output of gibbons, muriqui, chimpanzees, orangutan, bonobos and human vocalizations reveal some interesting features. Gibbons have long been recognized as producing elaborate and loud sounds. One interesting feature of these calls is that the sound source is predominantly produced by a whistled source (i.e. with an almost pure sinusoidal waveform (Figure 1)).

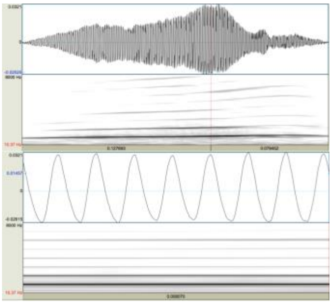


Figure 1. Audio waveform, narrowband spectrogram and a zoom of the waveform taken at the level of the cursor for a Gibbon whistled sound.

The fundamental frequency (F0) of these calls is often above 500 Hz and produces a spectrum with distant harmonics. This high F0 is intense and likely ideal for long distance propagation in dense rain forest. The best explanation so far for this whistled sound source is that it is produced between two stretched and non-oscillating vocal folds. The space between the arytenoids and the vocal folds acts as a kind of whistle. This type of whistled sound source is also observed among various monkeys and apes species such as muriqui, chimpanzees and bonobos. The comparative anatomy of gibbon, chimpanzee, bonobo and human show that the thickness of the vocal folds and the ratio between the respiratory glottis, made of the inter-arytenoid space, and the membranous glottis varies from one species to the other. This configuration of the glottis is found in other species where whistled sound sources are observed (it is even found with horses). Bonobos where the inter-arytenoid space occupies a large part of the glottis show an interesting feature that is the combination of a whistled and a low frequency sound source, i.e. a double source (Figure 2).

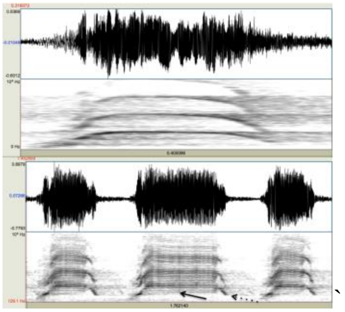


Figure 2. Audio waveform and narrowband spectrogram of a whistled Bonobo vocalization and three repetitions of a double source Bonobo vocalization. Arrows show the harmonics produced by the membranous glottis and by the whistled source between the inter-arytenoid space.

This is the combination of a whistled source produced in the inter-arytenoid space and the vibration of the membranous glottis situated in front of the arytenoid cartilages. These observations resulting from the comparison of anatomical data and the acoustics of non-human primate vocalizations suggest that one aspect the sound source's evolution in primates could have been the evolution from a whistled sound source with a high F0 to a lower frequency sound source. The change from whistled sources to lower frequency vibrations produced by oscillating vocal folds would eventually be the result of the increase in the size of the membranous glottis and the shortening of the arytenoids length. The lower frequency is also the result of a larger, thicker and less stiff membranous glottis (Harrison 1995). These anatomical changes were also accompanied by slight changes in the nerve supply of the larynx. These comparative data suggest that small changes in the anatomy and physiology of the vocals folds and glottis dimensions played an important role in the evolution of the sound source in primates.

References

Fitch, W.T., de Boer, B., Mathur, N. & Ghazanfar, A.A. (2016). Monkey vocal tracts are speech-ready. *Sci. Adv*. 2016; 2: e1600723.

Harrison, D.F.N. (1995). *The anatomy and physiology of the mammalian larynx*. Cambridge. Cambridge University Press.

Hirano, M. (1991. Phonosurgical anatomy of the larynx. In '*Phonosurgery, assessment and surgical management of voice disorders*'. C.N. Ford & D.M. Bless (eds.). Raven Press. New-York.

Starck, D. & Schneider, R. (1960). Respirationsorgane. In *Primatologia* III/2. H. Hofer, A.H. Schultz and D. Starck (eds.). S. Karger Basel.

Schön Ybarra, M. A. (1995). A comparative approach to the non-human primate vocal tract: implications for sound production. In E. Zimmerman & J.D. Newman & U. Jurgens (eds.) *Current topics in primate vocal communication*. Plenum Press. New-York.

IS THERE AN ICONIC MOTIVATION FOR THE MORPHOPHONOLOGICAL DISTINCTION BETWEEN NOUN-VERB PAIRS IN ASL?

KAREN EMMOREY*[1] and JENNIE PYERS[2]

*Corresponding Author: kemmorey@sdsu.edu
[1]Speech, Language and Hearing Sciences, San Diego State University, San Diego, USA
[2]Psychology Department, Wellesley College, Wellesley, USA

A match-to-sample, forced-choice paradigm was used in four experiments to test whether nonsigners are sensitive to the patterns of noun-verb derivation observed in many different sign languages. Nouns are often derived from verbs by (1) movement reduplication and size reduction or (2) size reduction alone (if the verb has reduplicated movement). Nonsigners (MTurk workers) were asked to match drawings of actions and objects with videos of noun and verb pairs in American Sign Language. Contrary to our predictions, nonsigners either were not sensitive to these rules (performing at chance) or they exhibited a bias that was the opposite of the one found in sign languages, preferring to match small reduplicated movements with actions, rather than with objects. Whether or not the noun-verb pairs exhibited pantomimic iconicity ("handling" signs or signs that resembled conventional pantomimes) did not affect the results. We speculate that internal linguistic pressures (a large lexicon, directional verbs) might override this weak, initial bias to map "more movement" (reduplication) to depict actions.

1. Introduction

Iconicity, typically defined as the resemblance of a symbol to its referent, is an essential tool in the creation of language, spoken or signed. Humans readily generate novel iconic symbols in the laboratory, and modern day languages bear traces of iconic origins. Some iconic mappings reflect associations between the linguistic form and common human cognitive construals of meaning (e.g. event telicity and sign movement; Strickland et al., 2015). However, the perception of iconicity can also be modulated by sign language knowledge (Occhino et al., 2017; Sehyr & Emmorey, 2019). We ask whether the morphophonological distinctions between nouns and verbs in American Sign Language (ASL) are iconically driven by cognitive construals of objects and actions. ASL nouns can be derived from verbs through two general principles: (1) movement reduplication + change in movement size (e.g., OPEN-BOOK vs. BOOK), or (2) change in movement size only if reduplication is already presented in the verb (e.g. PEDAL-BIKE vs. BIKE) (see Fig. 1; Supalla & Newport, 1978). One or both of these derivational strategies appear across many different sign languages (e.g., Italian Sign Language: Pizutto & Corazza, 1996; Russian Sign Language: Kimmelman, 2009; AUSLAN: Johnston, 2001; Israeli Sign Language: Tkachman & Sandler, 2013).

Across a set of four experiments, we tested non-signers' sensitivity to the two ways that nouns are derived from verbs. We hypothesized that these derivational strategies originated in universal human cognitive construals of actions and objects that can be iconically mapped to elements of the morphophonological structure of the sign (Wilcox, 2004). The clearest indicator of this construal would be the observation that nonsigners systematically associate large movements in signs with actions and small reduplicated movements with objects. We additionally hypothesized that pantomimic iconicity in a sign (see Fig. 1B) might affect non-signers assignment of noun-verb signs to objects and actions respectively.

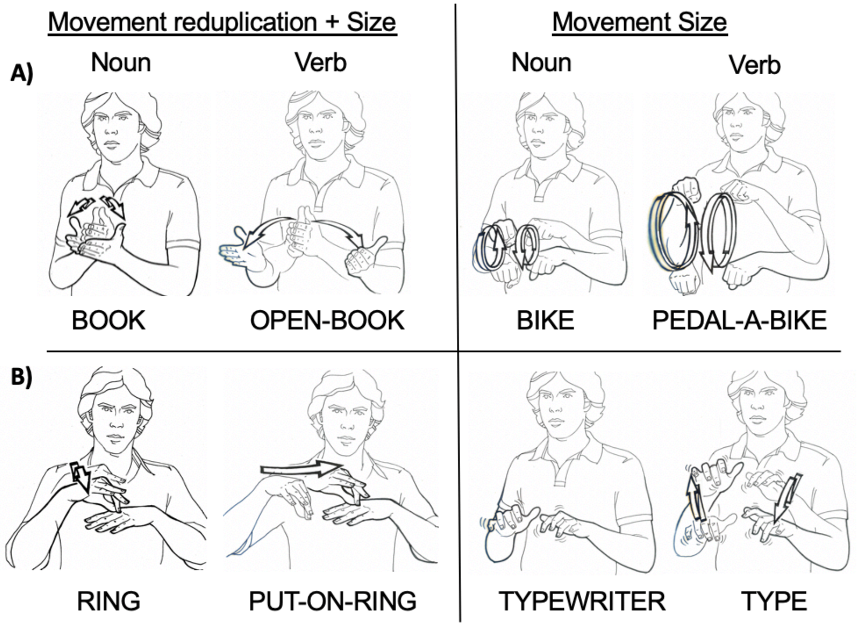


Figure 1. Examples of ASL noun-verb pairs derived by two morphophonological rules. **A)** Examples of non-pantomimic signs used in Experiments 1-3. **B)** Examples of pantomimic signs used in Experiment 4.

2. Methods

We selected a total of 61 noun-verb pairs in ASL that followed one of the rules shown in Fig. 1 and created line drawings that depicted the object and actions in each pair. For each experiment, we recruited 40 "master" MTurk workers who were non-signers. Participants viewed videos of the noun and verb signs and were asked to match an object and an action picture to each video (forced-choice matching task). We also coded whether signs were pantomimic and operationalized pantomimic signs as being handling signs (Fig. 1B) or conventional pantomimes that involve manipulating objects (e.g., holding a Y handshape for TELEPHONE/TO CALL).

3. Results

Exp. 1 assessed nonsigners' sensitivity to the *movement reduplication* + *size* rule for non-pantomimic signs (n = 20 noun/verb pairs). Participants' performance was significantly below chance (M=.223, SD=.196, $t(39)$=-8.935, p<.001). Nonsigners systematically associated movement reduplication with actions not objects, a pattern that is the opposite of what we see across sign languages.

Using the same forced-choice matching task, **Exp. 2** investigated whether non-signers show sensitivity to differences between ASL nouns and verbs that vary only in the *size of movement* (n = 9 pairs). Nonsigners exhibited some sensitivity to the expected pattern, mapping large movements to actions and smaller movements to nouns (M=.603, SD=.283, $t(39)$=2.305, p=.027).

Because these two derivational strategies co-exist in a sign language, **Exp. 3** presented nonsigners with a mix of 18 non-pantomimic noun-verb pairs (9 pairs following each rule). Non-signers were significantly below chance in making correct matches in both conditions, preferring to match small reduplicated movements to actions instead of objects (*Movement reduplication* + *size*: M= .291, SD=.258, $t(39)$=-5.123, p<.0001; *Movement size*: M=.391, SD=.247, $t(39)$=-2.80, p=0.008).

Exp. 4 tested whether these patterns held for pantomimic noun-verb pairs. For the *movement reduplication* + *size* rule (18 pairs), nonsigners again performed significantly below chance (M=.258, SD=.301, $t(40)$=-5.137, p<.001, systematically mapping small, reduplicated movements with actions. For the *movement size only* rule (15 pairs), nonsigners performed at chance, in contrast to Exp. 2 which included fewer noun-verb pairs (M=.455, SD=.341, $t(34)$=-.777, p=.443). When presented with a mix of rules (15 pantomimic noun-verb pairs following each rule), non-signers were at chance in making correct matches in both conditions (*Movement reduplication + size*: *M*=.416, *SD*=.328, *t*(38)=-1.569, *p*=.125; *Movement size only*: *M*=.522, *SD*=.264, *t*(38)=0.519, *p*=.607). However, non-signers were significantly more likely to attribute reduplication to actions over objects than to attribute a reduction in movement size to actions (*t*(38)=2.888, *p*=.006, 5%CI[0.031- 0.178]), even if they did so at a rate that was not greater than chance.

Finally, we conducted an exploratory analysis, using mixed effects logistic regression to understand the relative contributions of the type of rule and type of iconicity and the interaction between these two variables on mapping of the rule to the lexical class. Only rule type significantly predicted mapping changes in movement to changes in meaning (*B* =-0.75, *S.E.*= 0.28, *z*=-2.48, *p*=0.01). Participants exhibited more consistent mappings when the noun-verb contrast involved *movement reduplication + size* than with *movement size* alone. Although this consistent mapping was the inverse of the pattern observed in sign languages. Iconicity type and the interaction between rule type and iconicity type did not significantly predict accuracy (Iconicity type: *B* =-0.17, *S.E.*= 0.34, *z*=-0.49, *p*=0.63; Rule type x iconicity type: *B* =-0.00, *S.E.*= 0.39, *z*=-0,00, *p*=1.0).

3. Discussion

Across a set of four experiments we investigated whether morphophonological distinctions between nouns and verbs in ASL are iconically driven by the way humans construe objects and actions. We hypothesized that non-signers would be able to map ASL verbs to actions and nouns to objects, iconically mapping large and/or single movements to actions and small, reduplicated movements to objects. However, we found no evidence that non-signers followed the movement reduplication + size rule that is used cross-linguistically in sign languages to distinguish nouns and verbs. In addition, we found only modest evidence that non-signers followed the movement size only rule. Thus, nonsigners did not systematically associate large movements in signs with actions and small reduplicated movements with objects.

When viewing signs, nonsigners nevertheless did exhibit a set of iconic biases that differed from what is seen in sign languages. When forced to match the noun and verb variants to pictures of actions and objects, non-signers systematically interpreted movement reduplication and size reduction as more action-like and single path movements as more object-like at rates that differed significantly from chance. We speculate that nonsigners interpret reduplicated forms as having “more movement” and therefore mapped this form to actions. Goldin-Meadow et al. (1994) reported a similar pattern in a homesigner – David

– who systematically produced verb gestures with reduplication and noun gestures with a single movement (or fewer repetitions). Interestingly, the newest signers (Cohort 3) of Nicaraguan Sign Language were more likely to use movement reduplication to distinguish nouns from verbs compared to older NSL signers, suggesting that this noun-verb pattern is emergent (Abner et al., 2019). All NSL signers used movement size to distinguish nouns from verbs.

Given the robust evidence that nonsigners treat pantomimic iconicity as more action-like (e.g., Ortega & Özyürek, 2019), we speculated that we would see improved rule mapping with these signs. Pantomimic iconicity, however, did little to shift the biases we observed with signs that had no pantomimic iconicity. Indeed our exploratory regression analysis showed no effect of pantomimic iconicity when considering all of the data across the four experiments. Important to note, however, is that in our study handshape was not a contrastive feature for meaning and as such may have played little role in non-signers' attribution of meaning to the presented signs.

Taken together, we did not find systematic support for the idea that large articulated movements in a manual signal emphasizes the process oriented nature of actions (Wilcox, 2004) in a way that allows for nonsigners' systematic interpretation of large movements as actions and small movements as objects. Nor did we find that nonsigners interpret small reduplicated manual movement as de-emphasizing the action which in turn should favor an object interpretation of such movement (Lepic & Padden, 2017).

If nonsigners (and possibly homesigners) exhibit a bias to interpret "more movement" (i.e., reduplication) as iconically mapping to actions over objects, then why do sign languages systematically exhibit the opposite pattern? We speculate that this relatively weak cognitive construal effect could be over-ridden by internal linguistic pressures as a sign language develops. For example, with a large lexicon, small movements may be used to focus attention on the hands (which map to objects), and such small movements may need to be reduplicated for saliency. Supporting this speculation, recent work indicates that the movement of nouns is smaller than verbs across the ASL lexicon (Sehyr, Edinger, & Emmorey, 2019). In addition, path movements for verbs can be articulated directionally to specify changes in location or grammatical roles. Future research is needed to investigate whether such linguistic pressures override an initial iconic bias for construing reduplicated movement as associated with actions over objects.

Acknowledgements

This work was supported in part by a grant from the National Institute of Deafness and other Communication Disorders (DC010997) and by a grant from the James S. McDonnell Foundation. We wish to thank Aislinn Diaz and Hailey den Elzen for help in running this study.

References

Abner, N., Flaherty, M., Stangl, K., Coppola, M., Brentari, D., & Goldin-Meadow, S. (2019). The noun-verb distinction in established and emergent sign systems. *Language*, *95*(2), 230-267.

Goldin-Meadow, S., Butcher, C., Mylander, C., & Dodge, M. (1994). Nouns and verbs in a self-styled gesture system. What's in a name? *Cognitive Psychology, 27,* 259-319.

Johnston, T. (2001). The lexical database of AUSLAN (Australian Sign Language). *Sign Language & Linguistics*, *4*(1), 145-169.

Kimmelman, V. (2009). Parts of speech in Russian Sign Language: The role of iconicity and economy. *Sign Language & Linguistics*, *12*(2), 161-186.

Lepic, R., & Padden, C. (2017). A-morphous iconicity. In C. Bowern, L., Horn, & R. Zanuttini (Eds.), *On looking into words (and beyond): Structures, relations, analyses* (pp. 489-516), Berlin, Germany: Language Sciences Press.

Occhino, C., Anible, B., Wilkinson, E., & Morford, J. P. (2017). Iconicity is in the eye of the beholder. *Gesture*, *16*(1), 100-126.

Ortega, G., & Özyürek, A. (2019). Systematic mappings between semantic categories and types of iconic representations in the manual modality: a normed database of silent gesture. *Behavior and Research Methods.*

Pizzuto, E., & Corazza, S. (1996). Noun morphology in Italian Sign language (LIS). *Lingua*, *98*(1-3), 169-196.

Sehyr, Z., S., Edinger, R., & Emmorey, K. (2019). Quantifying differences in spatial and temporal patterns between nouns and verbs in American Sign Language. Poster presented at the Theoretical Issues in Sign Language Research meeting, September, Hamburg, Germany.

Sehyr, Z. S., & Emmorey, K. (2019). The perceived mapping between form and meaning in American Sign Language depends on linguistic knowledge and task: evidence from iconicity and transparency judgments. *Language and Cognition*, *11*(2), 208-234.

Supalla, T, & Newport, E. (1978). How many seats in a chair? In P. Siple (Ed.), *Understanding language through sign language research* (pp. 158-181). New York: Academic Press.

Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., & Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences*, *112*(19), 5968-5973.

Tkachman, O., & Sandler, W. (2013). The noun–verb distinction in two young sign languages. *Gesture*, *13*(3), 253-286.

Wilcox, S. (2004). Cognitive iconicity: Conceptual spaces, meaning, and gesture in signed language. *Cognitive linguistics*, *15*(2), 119-148.

OPEN COMPOSITIONALITY IN PIED BABBLER CALL COMBINATIONS

SABRINA ENGESSER*[1,2], AMANDA R. RIDLEY[3,4], STUART K. WATSON[1,2],
SOTARO KITA[5], and SIMON W. TOWNSEND[1,2,5]

*Corresponding Author: sabrina.engesser@outlook.com; Current Address: Department of Cognitive Biology, University of Vienna, 1090 Vienna, Austria

[1]Center for the Interdisciplinary Study of Language Evolution, University of Zurich, 8032 Zurich, Switzerland

[2]Department of Comparative Linguistics, University of Zurich, 8032 Zurich, Switzerland

[3]Centre for Evolutionary Biology, School of Animal Biology, The University of Western Australia, Crawley WA 6009, Australia

[4]Percy FitzPatrick Institute, University of Cape Town, Rondebosch 7701, South Africa

[5]Department of Psychology, University of Warwick, Coventry CV4 7AL, United Kingdom

Human language's generative power rests largely on its compositional nature: words can be assembled into novel, perfectly intelligible, higher-order structures with derived meaning (Hurford, 2012). The generativity of such a compositional system depends on how "open" it is to modifications, namely whether compositional structures can be modified (e.g. changes to word order) but still remain comprehensible to listeners. For example, if one understands the compositional meaning of "*Danger ahead!*", one can also interpret the uncommon form "*Ahead danger!*" (see Goldberg (1995) for related examples). Whilst "open compositionality" is ubiquitous in human language, whether animals show any precursor to such a capacity is controversial, yet key to understanding the evolution of language's generativity (Bolhuis, Beckers, Huybregts, Berwick, & Everaert, 2018; Townsend, Engesser, Stoll, Zuberbühler, & Bickel, 2018).

Here we provide evidence that a non-human animal is capable of deriving the meaning of a structurally modified, species-own compositional structure. Pied babblers (*Turdoides bicolor*), a highly social bird species from Southern Africa, produce alert-recruitment call combinations (Fig. 1) when encountering predators, eliciting mobbing behaviour with prolonged alertness and instantaneous recruitment of conspecifics to the threat (Engesser, Ridley, &

Townsend, 2016). Previous research indicated this call complex to be compositional, with the meaning of the whole being related to, yet going beyond, the meanings of its component calls (*s. l.* "approach to/because of a threat", as opposed to "be alert and approach"). Through conducting behavioural observations and predator presentation experiments we here show that the combination is highly conserved in its structure with the alert call always preceding the recruitment call component when combined. Despite this stereotyped structure, however, we demonstrate that modifications to the call combination can nonetheless be comprehended by receivers. Specifically, playback experiments exposing pied babblers to an unnatural, reversed call order of their alert-recruitment combination (i.e. recruitment-alert combination, Fig. 1) revealed that adult birds responded in identical ways to natural and reversed-ordered variants. These data suggest that the alert-recruitment call combination represents a compositionally open structure that can tolerate modifications, likely through babblers' ability to relate the meaning of the whole to its comprising, individually meaningful calls. Furthermore, we found that younger individuals responded less to the reversed-ordered variant, suggesting that a "compositional sensitivity" might be experience-based and acquired during ontogeny, with call combinations initially being perceived holistically as unsegmented chunks.

To conclude, we provide evidence for open compositionality in the communication system of a non-human animal species: despite being stereotyped in production, the babbler combinatorial system is compositionally open in perception enabling the processing of novel call combinations. Such compositional sensitivity might therefore represent a key prerequisite for generativity and a key stepping-stone on the evolutionary path to language's full blown infinite compositional productivity.

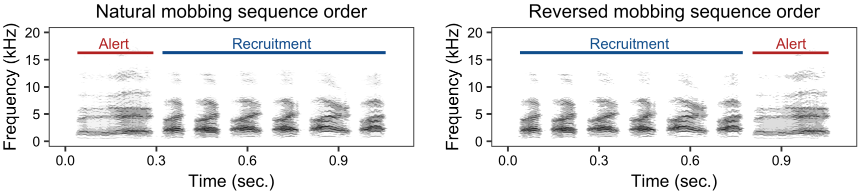


Figure 1. Spectrograms of a natural (alert-recruitment) and a reversed-ordered (recruitment-alert) call combination. Alert calls are typically composed of 1-2 call elements, recruitment calls of 4-7 call elements.

References

Bolhuis, J. J., Beckers, G. J. L., Huybregts, M. A. C., Berwick, R. C., & Everaert, M. B. H. (2018). Meaningful syntactic structure in songbird vocalizations? *PloS Biology, 16*, e2005157.

Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of Sciences, 113*, 5976-5981.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Hurford, J. (2012). *The origins of grammar*. Oxford: Oxford University Press.

Townsend, S. W., Engesser, S., Stoll, S., Zuberbühler, K., & Bickel, B. (2018). Compositionality in animals and humans. *PLoS Biology, 18*, e2006425.

COMMUNITY STRUCTURE AFFECTS CONVENERGENCE ON UNIFORM WORD ORDERS: EVIDENCE FROM EMERGING SIGN LANGUAGES

RABIA ERGIN*[1], LIMOR RAVIV[2], ANN SENGHAS[3], CAROL PADDEN[4], and WENDY SANDLER[5]

*Corresponding Author: rabia.ergin@gmail.com
[1]Cognitive Science Program, Boğaziçi University, Istanbul, Turkey
[2]Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
[3]Department of Psychology, Barnard College, New York City, USA
[4]Department of Communication, University of California, San Diego, USA
[5]Department of English Language and Literature, University of Haifa, Haifa, Israel

Community Structure and Emerging Sign Languages

Previous research suggests that the size and social structure of the community may play a crucial role in shaping linguistic features in the early stages of language evolution (e.g., Lupyan & Dale, 2010; Meir & Sandler, to appear; Raviv, Meyer, & Lev-Ari, 2019). Specifically, it was hypothesized that emerging sign languages that develop in small and tightly-knit communities (village sign languages) tend to have less conventionalized structure. In contrast, languages emerging in bigger communities and/or in communities whose members have less shared background (deaf community sign languages) tend to be more uniform (Meir, Israel, Sandler, Padden, & Aronoff, 2012).

Here, we directly investigate this hypothesis by focusing on a fundamental feature of languages: word order. Word order is one of the most basic linguistic conventions, and is essential for expressing semantic roles (i.e., who did what to whom). Using naturalistic data, we aim to uncover how such a key element becomes a convention in new communication systems that differ in their social structure and in their community size.

To this end, we provide the first direct comparison of word order variability (a proxy for the degree of conventionalization) in three different emerging sign languages: (1) Central Taurus Sign Language (CTSL), a village sign language used in a small community Turkey (e.g., Ergin, 2017); (2) Al-Sayyid Bedouin Sign Language (ABSL), a bigger village sign language used in Israel (e.g., Sandler, Meir, Padden, & Aronoff, 2005); and (3) Nicaraguan Sign Language (NSL), an even bigger deaf community sign language used in Nicaragua (e.g., Senghas, 1995). Given that these three languages differ in community size and social structure, they allow us to investigate whether and how these properties influence the formation of word order conventions in the early stages of language evolution.

The Current Study

We examined production data from 14 CTSL and 25 ABSL deaf signers. These two communities share certain social characteristics but differ in community size. NSL represents a still larger community with different social characteristics, and data coding and analyses for NSL are in progress. All participants watched 18 short video clips, originally developed by Sandler et al. (2005). Signers were asked to describe each clip to a deaf/hearing addressee, resulting in 381 elicited responses from CTSL and 543 from ABSL signers. The clips included three types of scenarios: six transitive irreversible events in which a human agent acts on an inanimate patient (e.g., MAN-WATERMELON-TAP), six transitive reversible events in which a human agent acts on a human patient (e.g., MAN-GIRL-TAP) and six ditransitive events in which two human characters are involved in a transfer relation (e.g., MAN-GIRL-BALL-THROW).

We coded signers' productions for word order using a comparable scheme, and examined the degree of variability within signers (i.e., whether a single signer is consistent in their selected word order in a given scenario), and across signers (i.e., indicating the degree of conventionalization in the community).

Results

Preliminary results show that in all three types of scenarios, there is significantly more variance in word order preferences in CTSL as opposed to those in ABSL, both within and across signers: CTSL signers show less convergence as a community (i.e., producing significantly more word order variants across different signers, p=0.016), and are less consistent in their own productions (i.e., producing significantly more word order variants as individuals, p=0.013). These results support the hypothesis that the size of a language community has an effect on conventionalization in early stages of language emergence: the language of bigger communities is more uniform in structure that that of smaller communities (see Figures 1-2).

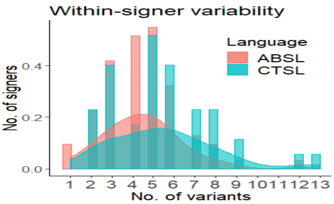


Figure 1. Within-signer variability

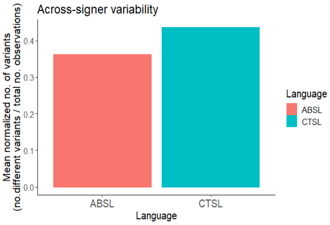


Figure 2. Variability across signers (convergence)

References

Ergin, R. (2017). *Central Taurus Sign Language: A Unique Vantage Point into Language Emergence.* (Ph.D. Ph.D. Thesis), Tufts University Retrieved from https://search.proquest.com/openview/69f73091abbc97d1d2997408076c5301/1?pq-origsite=gscholar&cbl=18750&diss=y

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PloS one, 5*(1), e8559.

Meir, I., Israel, A., Sandler, W., Padden, C., & Aronoff, M. (2012). The influence of community on language structure: evidence from two young sign languages. *Linguistic Variation, 12*(2), 247-291.

Meir, I., & Sandler, W. (to appear). Variation and conventionalization in language emergence: A study of two young sign languages of Israel. In E. Doron, M. Rapapaort Hovav, Y. Reshef, & M. Taube (Eds.), *Lingusitic Contact, Continuity and Change in the Genesis of Modern Hebrew*. Amsterdam: John Benjamins.

Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Larger communities create more systematic languages. *Proceedings of the Royal Society B, 286*(1907), 20191262.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America, 102*(7), 2661-2665.

Senghas, A. (1995). *Children's contribution to the birth of Nicaraguan Sign Language.* Massachusetts Institute of Technology.

LINGUISTIC COMPLEXITY INCREASES AS A FUNCTION OF ATTRACTIVENESS IN INTERSEXUAL COMMUNICATION. TENTATIVE EXPERIMENTAL SUPPORT

CHARLOTTE ESSERS, FREEK VAN DE VELDE[*]

[*]freek.vandevelde@kuleuven.be
Dept. of Linguistics – QLVL, KU Leuven, Belgium

Language is not merely a tool for transmitting information, but serves other functions as well, though there is neither a shortage of theories, from language's role in social grooming (Dunbar 1996) to the Beau-Geste hypothesis about territorial defense (Fitch 2000: 264), nor a lack of controversy surrounding these theories. One of the more robust ideas is that language plays a crucial role in courting behavior. Vocalization is especially important in tournament avian species, and though humans are arguably less tournament-oriented (Prum 2017), they still have various ways to signal sexual fitness. Given the highly communicative nature of human behavior, it comes as no surprise that language plays an important role here, both as a signal of fitness by ostentatious display of a costly trait (Miller 2000: 305; Rosenberg & Tunney 2008), as a runaway process in sexual antagonism to detect deception (Ridley 1999: 115-116), and in the context of assortative mating. The idea that language evolved through sexual selection in fact goes back to Darwin (1871) already.

In this paper, we set out to test whether we could detect an impact of courtship behavior on language. If language is indeed a signal of fitness, one may expect that language complexity will increase during flirtation. On the other hand, there is research showing that men may be cognitively impaired when talking to attractive women (Karremans et al. 2009), a finding that is potentially at variance with the drive to display higher complexity.

For ethical as well as practical reasons, it is hard to induce situations of flirtation in a controlled experiment, but in reality, similar behavior can already be detected in normal interaction. Earlier research has shown that men show different behavior depending on whether they find the interviewer attractive or not. In the classical ‘frizzy wig’ experiment by Sigall & Aronson (1969), for instance, male test subjects turned out to be more sensitive to the results of a personality test when it was administered by an attractive woman than by an unattractive woman. Starting from this observation, we conducted an experiment to see whether the attractiveness of the interviewer had an impact on the interviewee’s language complexity.

The experiment was conducted with 40 male participants between the ages of 18 and 25 years old. As part of the between-subject design, 20 participants were interviewed by a confederate in an attractive condition, while the other 20 participants were interviewed by the same confederate in a non-attractive condition. Participants were not aware of the experiment’s aim to avoid the observer’s paradox. As a distractor, participants were asked to find an optical illusion in ten photographs and were later asked to answer some personal questions about their family, education and personality. The difference in attractiveness of the interviewer was emphasized through make-up, glasses and clothing. To make sure the difference in perceived attractiveness was real, subjects were asked to rate the interviewer’s attractiveness, write down their answer anonymously on a slip of paper, which was then put in an urn (per condition). The interviews were transcribed verbatim and subjected to linguistic analysis by the application T-Scan, which measures linguistic complexity on several predictors for Dutch. Of these predictors, the following were selected for their intuitive correlation with linguistic complexity: total number of words, total number of sentences, morphemes per word, word frequency, sentence length, level of syntactic depth, type-token-ratio, lexical diversity, and the use of connectives (Kraf & Pander Maat 2009; Pander Maat et al. 2017). As we need the whole verbatim transcript per participant for the calculation of the complexity metrics, the total number of observations per condition is limited (2 x 20), precluding an analysis in which the different metrics can be entered as independent variables in a logistic regression with the attractiveness condition as the outcome. Such an analysis would also suffer from multicollinearity. Instead, we carried out one t-test per complexity metric, treating the attractiveness as the independent variable. Even though the difference in the stimulus is fairly small, with just two different appearances of the same woman, the results reveal that the interviewees' language was consistently higher in the attractive condition, though the results do not reach the common alpha level of 0.05, due to the restricted size of the participants sample, leading to underpowered tests. Still, the consistency of the difference in all the tests applied suggest – albeit tentatively – that Darwin's sexual selection theory of language still leave a detectable signal in present day language use.

References

Darwin, C. 1871. *The descent of man, and selection in relation to sex*. London: John Murray.

Dunbar, R. 1996. *Evolution of language*. London: Faber & Faber.

Fitch, W. T. 2000. 'The evolution of speech: a comparative review'. *Trends in Cognitive Sciences* 4(7): 258-267.

Karremans, J., T. Verwijmeren, T. Pronk. & M. Reitsma. 2009. 'Interacting with women can impair men's cognitive functioning'. *Journal of Experimental Social Psychology* 45(4): 1041-1044.

Kraf, Rogier & Pander Maat, Henk. 2009. 'Leesbaarheidsonderzoek: oude problemen, nieuwe kansen'. *Tijdschrift voor Taalbeheersing* 31(2): 97-123.

Miller, G. 2000. *The mating mind. How sexual choice shaped the evolution of human nature*. New York: Doubleday.

Pander Maat, H., R. Kraf & N. Dekker. 2017. *Handleiding T-Scan*.

Prum, R.O. 2017. *The evolution of beauty. How Darwin's forgotten theory of mate choice shapes the animal world – and us*. New York: Doubleday.

Ridley, M. 1999. *Genome. The autobiography of a species in 23 chapters*. London: Fourth Estate.

Rosenberg, J. & R. Tunney. 2008. 'Human vocabulary use as display'. *Evolutionary Psychology* 6(3): 538-549.

Sigall, H. & E. Aronson. 1969. 'Liking for an evaluator as a function of her physical attractiveness and nature of the evaluations'. *Journal of Experimental Social Psychology* 5(1): 93-100.

COLOR SALIENCE, COLOR TERM EVOLUTION, AND THE ISSUE OF RELATEDNESS

CALEB EVERETT *1, COLLEEN O'BRIEN 2

* Corresponding Author: caleb@miami.edu
1Anthropology, University of Miami, Coral Gables, USA
2 JCRS, Friedrich Schiller Universität, Jena

Gibson et al. (2017, GEA henceforth) calculated the average surprisal or "entropy" of languages' basic color terms, based on the data of the World Color Survey (Kay et al. 2009). Their innovative work suggests, *inter alia*, that industrialization may yield less entropy and greater efficiency of basic color terms. They obtained the average "surprisal" of each color chip named in the survey, for each language. In their words, "The surprisal score for each color *c* is computed by summing together a score for each word *w* that might have been used to label *c*, which is calculated by multiplying P(w|c) by -log(P(c|w)), the listener's surprisal that *w* would label *c*. We estimate P(c|w) via Bayes Theorem assuming a uniform prior on P(c)." Formulaically:

$$S(c) = \sum_{w} P(w|c) \log \frac{1}{P(c|w)}$$

Via this approach, GEA find that "warmer" Munsell colors (e.g. red and yellow) are encoded with greater efficiency or less surprisal, when contrasted with "cool" colors (e.g. green and brown). This central finding holds across the 110 languages in the WCS plus 3 others tested. The approach can also be used to generate an average "entropy" score for each of these 113 languages by averaging all the color chips' surprisal values in a given language.

We used GEA's code to ascertain languages' average entropy scores, to more carefully examine a secondary claim made in their work: "industrialization… increases color usefulness" and helps to reduce the entropy of color terms in a language. (10785) Color terms may evolve to more efficiently encode the hues associated with foregrounded objects in industrialized cultures, objects which tend to be warm-colored according to GEA's ancillary analysis of thousands of images. GEA note that English and Spanish have the most efficient color term sets in their data, consistent with the industrialization hypothesis. Yet our reanalysis of the WCS data suggests no clear association between industrialization and color terms, given the confound of language relatedness. The entropy scores (average per-color surprisal) of languages range from 3.08 (English) to 5.88 (Eastern Cree) (mean=4.7, s.d.=0.37, n=113). The most "efficient" languages consist primarily of two clusters that deviate from typical efficiency values: a) Indo-European languages and European-based creoles and b) Mesoamerican languages. At the other end of the efficiency/entropy spectrum, Algonquian languages tend to have high entropy color terms. See Figure 1 for a visual summary of the results across all families. In short, a few language families and one geographic region are over-represented at the two ends of the color-entropy continuum. Indo-European languages could very well have low entropy colors because of industrialization, but we conclude that any claims regarding the causal role of industrialization are difficult to support with the WCS data. The confound of language relatedness, not to mention language contact, cannot be adequately controlled for with WCS-based efficiency data.

As the exception that proves the rule, we describe the case of the Colombian isolate Kamsá. Kamsá has a surprisingly efficient color term system, superficially violating the expectations of GEA's secondary hypothesis. Based on our (CO) own research with Kamsá, however, we discuss how the WCS data were gathered in part or entirely with bilingual Kamsá speakers, given the high rates of Spanish bilingualism since the 1970's. (With bilingualism in Inga prevalent prior to Spanish bilingualism.) We suggest that the efficiency of Kamsá responses may be the result of that Spanish bilingualism, further pointing towards relatedness as a confound in the color-term efficiency findings. It seems quite plausible that industrialization has impacted color terms in the ways GEA suggest. Yet their secondary hypothesis cannot be adequately supported by the WCS data, given the potential role of common ancestry in shaping color term entropy.

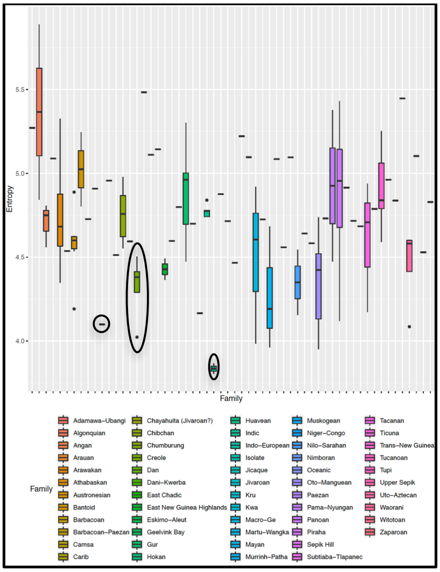


Figure 1. Box plots of color-term entropy scores in the WCS, by language family. Kamsá, creoles, and IE languages (in that order, from left to right) are highlighted.

References

Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., Gibson, M., Piantadosi, S. & Conway, B. (2017). Color naming across languages reflects color use. *PNAS 114*, 10785-10790.

Kay, P., Berlin, B., Maffi, L., Merrifield, W. & Cook, R. (2009) *The World Color Survey.* Stanford: CSLI.

EVOLUTIONARY MECHANISMS IN THE BENGALESE FINCH'S SONG: PARALLELS AND IMPLICATIONS FOR THE STUDY OF HUMAN SPEECH

MADZA FARIAS-VIRGENS*[1], TERRENCE DEACON[2], KAZUO OKANOYA[3],
STEPHANIE WHITE[4], EMILIA HUERTA-SANCHEZ[5]

*Corresponding Author: madzayasodara@ucla.edu
[1]Interdepartmental Graduate Program in Molecular, Cellular and Integrative Physiology, University of California Los Angeles, Los Angeles, United States of America
[2]Department of Anthropology, University of California Berkeley, Berkeley, United States of America
[3] Graduate School of Arts and Sciences, The University of Tokyo & RIKEN-Brain Science Institute, Tokyo, Japan
[4] Department of Integrative Biology and Physiology, University of California Los Angeles, Los Angeles, United States of America
[5]Department of Ecology and Evolutionary Biology, Brown University, Providence, United States of America

A robust body of evidence accrued over >100 years of research in birdsong demonstrates striking analogies between this exquisite behavior practiced by a much evolutionarily distant animal group and speech in our species. Both birdsong and speech depend upon vocal production learning, that is the ability to learn how to produce vocal sounds via imitation - in addition to vocal usage learning, the ability to associate innate sounds with specific elements and events. Like a child learning to speak, a young bird must first hear the vocal sounds of adults and then imitate those sounds on its own. Initially, the bird will sing a faint, unstructured song, akin to babbling in human infants. By adulthood, this immature chirping progresses to a more consistent birdsong, just as babbling progresses to speaking. These parallel developmental trajectories are accompanied by striking similarities in the way brains and their underlying molecular apparatus functions to produce speech and birdsong. All this resemblance has motivated the additional search for similar evolutionary pressures leading to vocal learning in songbirds and humans. Our research uses a songbird system to identify evolutionary processes leading to increased complexity of learned vocal behavior, a key aspect also in speech evolution. The Bengalese finch (BF) has a remarkably complex song, in which transitions between vocal units are less firmly fixed, introducing variability in song sequencing. This vocal complexity evolved during BF's domestication from the white-backed munia (WBM). Our study investigates genetic signatures left by evolutionary forces shaping BF's song complexity. For this, we have sequenced whole genomes of individuals within the two bird strains (11 BFs and 11 WBMs). Using analytical tools from comparative and population genomics, we have been able to access the relative contributions of selection processes, such as female choice for more complex songs, and relaxation of sources of evolutionary constraints to song complexity that are commonly found in the wild but absent in the domesticated scenario, such as stress related to finding food or defending from predators or pressures to avoid confusion with cohabitating finch species. We have identified a group of about 300 out of 14.000 genes highly differentiated between the two bird strains, as evidenced by a higher fixation index (Fst), which indicates greater variability between relative to within populations. This list includes genes with a known role in bird song, such as DRD3, which encodes the dopamine D3 receptor and has also been linked to schizophrenia, as well as genes causally related to important motor and sensory neuropathies, such as HK1 which codifies the brain form hexokinase and when mutated leads to Charcot-MarieTooth disease. We also introduce sets of genes for which observed variability deviates from the expected in each bird strain, as evidenced by Tajima's D. We report overall reduced levels of genetic variability in the BF. However, genes showing increased variability in the BF relative to its ancestor strains also exist. We have inferred a demographic model shaping BF's genetic variation, which corroborates historical reports, and have estimated the impact of the population bottleneck during domestication. Our results have the potential to guide further comparative efforts toward identifying similar patterns of evolutionary change between humans and other primates or hominid lineages (e.g. Denisovans and Neanderthals).

FUNDAMENTAL FLAWS IN MINIMALIST ACCOUNTS OF LANGUAGE: REMOVING BARRIERS TO AN UNDERSTANDING OF LANGUAGE EVOLUTION

ANDREW FEENEY

*Corresponding Author: andrew.feeney@northumbria.ac.uk
Department of Humanities, Northumbria University, Newcastle upon Tyne

This year sees the 25th anniversary of the publication of Chomsky's 'the Minimalist Program' (1995), and while generative linguistics now faces multiple challenges to its philosophical basis, as well as its concomitant linguistic analyses – most notably from the broad school characterized as *Cognitive Linguistics* (e.g. Geeraerts & Cuyckens, 2007) – the generative approach maintains a significant influence in the field of cognitive science. And while in the early days, generative linguistics was both motivated by a need to account for Plato's problem (the logical problem of language acquisition by which children appeared able to acquire a linguistic system in a rapid and systematic manner despite the impoverished nature of their primary linguistic data) and had a significant impact on the course of the study of language acquisition, so more recently (if to a lesser extent) generative linguistic accounts of the human faculty for language have had to be defended in terms of a plausible evolutionary scenario, and have consequently had an impact on the nature of evolutionary linguistic inquiry.

However, Minimalism today is not a single unified theory of language. There is, for example, a substantial difference in the understanding of the properties of lexical items and the operations that underpin derivation. On the one hand there are those (including Chomsky) who adhere to a *Strong Minimalist Thesis* in which the core linguistic operations are reduced to an absolute minimum (perhaps only labelling and merge) whereby undesirable derivations are filtered out at the non-linguistic interfaces. In contrast there are minimalists who argue for a complex system of features which constrain merge in such a way that non-grammatical derivations are avoided – so called *crash-proof* grammars (see Putnam, 2010). Roughly corresponding to each of these perspectives respectively are proponents of a rapid emergence of complex language (Berwick & Chomsky, 2016; Berwick & Chomsky, 2017) and those who propose a classical Darwinian gradualism (e.g. Pinker & Jackendoff, 2005; Jackendoff &

Pinker, 2005). I will argue that both explanations are fundamentally flawed in their conceptions of the human faculty for language and, more significantly, their account of its evolutionary origins.

An examination of the evidence from contributory disciplines to the study of hominin evolution reveals two clearly distinguishable periods in the evolution of human cognition corresponding to stages of ‘punctuated equilibrium’ (Gould & Eldridge, 1993). The first of these followed a period, approximately two million years ago, in which there were a comparatively large number of changes to genes and genomic regions (especially in the Human Accelerated Region 1) with consequences for brain lateralization, connectivity and overall brain size (Kamm et al., 2013). Newly emerging behaviours evidenced at this time, and indicative of significant cognitive changes, include sophisticated (mode 2) tool use, migration out of Africa and probable use of fire. Such accomplishments are indicative of cooperation beyond that exhibited by any preceding hominins (Tomasello et al., 2012).

While it is possible that these premodern humans were communicating in a more sophisticated way than any other species had achieved at this time, via a form of a simple protolanguage of concatenated symbols, the following period of a million years or so was characterized by almost complete cultural stasis indicative of a species with a cognitive capacity that lacked substantial creativity. In ‘Dual Processing Theory’ (DPT) this is understood as ‘System’ 1, which is evolutionary ancient and shared with other species (see Evans & Frankish, 2009). A second intense episode of evolutionary activity followed around 500 thousand years ago involving further growth and reorganization of the brain which resulted in an advanced theory of mind and a uniquely human, hierarchically structured, creative System 2 type cognitive processing. Extant systems for intentional communication were coopted for the external representation of this increasing complex mind-internal cognition, becoming increasingly complex as the task required.

This account provides further support to the notion of language in the ‘Representational Hypothesis’ (e.g. Burton-Roberts, 2011) in which we posit a mind internal structured cognition – the sole locus of semantics and syntax – and a distinct phonological system of representation. I argue that this mind internal, hierarchically structured thought is not realized externally as in the Minimalist Program, but is rather represented in language in the form of a symbolic, semiotic system grounded purely in the properties of the articulatory-phonetic system. There is no hierarchically structured syntax in language nor are there any semantic properties. Language is a complex system of linear phonological representation subject to human pragmatic interpretation. Adherence to an orthodox Minimalist notion of language, the search for the cognitive foundations of linguistic syntax and semantics (e.g. Friederici et al., 2017) and their evolutionary origins, will only impede progress in the understanding of language evolution.

References

Berwick R. C. & Chomsky, N. (2016) *Why only us*. Cambridge, MA.: MIT Press

Berwick R. C. & Chomsky, N. (2017) Why only us: Recent questions and answers. *Journal of Neurolinguistics*. 43 166-177

Burton-Roberts, N. (2011). On the grounding of syntax and the role of phonology in human cognition. *Lingua*. 121, 2089-2102;

Chomksy, N. (1995) The Minimalist Program. Cambridge, MA.: MIT Press

Evans, J. S. B. T. & Frankish, K. (eds.) (2009). In two minds : dual processes and beyond. Oxford: Oxford University Press

Frediederici,D. A., Chomsky, N., Berwick, R. C., Moro, A., & Bolhuis, J. (2017) Language, mind and brain. *Nature Human Behaviour* 1, 713–722

Geeraerts, D. and H. Cuyckens (ed.) (2007). *The Oxford Handbook of Cognitive Linguistics*.Oxford: Oxford University Press

Gould, S. J. & Eldridge, N. (1993) Punctuated equilibrium comes of age. *Nature*. 366, 223-227

Jackendoff, R. & Pinker, S. (2005) The nature of the language faculty and its implications for evolution of language. *Cognition*. 97, 211-225

Kamm, G. B., Pisciottano, F. Kli ger, R. & Franchini, L. F. (2013). The Developmental Brain Gene NPAS3 Contains the Largest Number of Accelerated Regulatory Sequences in the Human Genome. *Molecular Biology and Evolution*. 30, 1088-1102

Pinker, S. & Jackendoff, R. (2005) The faculty of language: what's special about it? *Cognition*. 95, 201-236

Putnam, M. T. (ed.) (2010) *Exploring Crash-Proof Grammars*. Amsterdam: John Benjamins.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E. & Herrmann, E. (2012) Two Key Steps in the Evolution of Human Cooperation: The Interdependence Hypothesis. Current Anthropology. 53, (6) 673-692

PHONEME INVENTORY SIZE AND THE LANGUAGE SYSTEM

GERTRAUD FENK-OCZLON[*1]

[*] Corresponding Author: gertraud.fenk@aau.at
[1] IFK, Alpen-Adria Universität, Klagenfurt, Austria

Phoneme inventory size and its relationship with linguistic and non-linguistic factors remains a matter of debate. Especially research reporting correlations between inventory size and non-linguistic factors like population size (Hay & Bauer, 2007) or distance of Africa (Atkinson, 2011) gained a lot of interest. Concerning relationships within the language system Maddieson (2006) and Fenk-Oczlon & Fenk (2008) found a positive correlation between inventory size and syllable complexity; Nettle (1995, 1998) and Wichmann et al. (2011) report an inverse relationship between word length and inventory size. All these studies use single words for their correlations or descriptions of the permitted syllable structures in the respective languages. But the length of uninflected words in dictionaries or word lists or the permitted maximum syllable complexity in individual languages do not reflect word length or syllable size in textual material or actual language use (cf. Maddieson 2009).

A new study: Cross-linguistic correlations between phoneme inventory size, syllable size, and word length in textual material

Hypothesis: Cross-linguistic correlations between phoneme inventory size, syllable size and word length show also in textual material. ***Method:*** Parallel texts were obtained by an elicitation experiment: Native speakers of 56 languages from 17 language families were asked to translate a matched set of 22 simple declarative sentences into their mother tongue. We then calculated the average syllable complexity (in number of phonemes) and the average word length (in number of syllables and number of phonemes) in these texts and correlated the data with the size of the language's phoneme inventories found in UPSID and/or the PHOIBLE database. ***Results:*** • a significant positive correlation between syllable complexity and phoneme inventory (r = .58, n = 56, p < .01*)*; • a significant negative correlation between word length in number of syllables and phoneme inventory (r = –.38, n = 56, p < .01); • an insignificant negative correlation between phoneme inventory size and number of phonemes per word (r = –.14, n = 56, ns.**)** ***Discussion****:* A significant positive correlation between inventory size and syllable complexity as well as a significant negative correlation between inventory size and word length in number of syllables shows also in texts. However, Nettl's and Wichmann et al.'s negative correlation between inventory size and word length in number of phonemes turned out to be insignificant in our textual material. This rather unexpected result will be explained by referring to a cross-linguistic version of Menzerath's law (Fenk-Oczlon & Fenk, 1999).

Phoneme inventory size: interactions with the language system

The present study demonstrates that phoneme inventory size is strongly correlated with syllable complexity in number of phonemes. Syllable complexity in turn is associated with short words and morphological simplicity, as shown in previous studies (Fenk-Oczlon & Fenk, 1999, 2005; see Tab.1)

Table 1. Associations between syllable complexity, phoneme inventory and other linguistic features

high syllable complexity	low syllable complexity
large phoneme inventory size	**small phoneme inventory size**
low number of syllables per word	high number of syllables per word
high number of monosyllables	low number of monosyllables
low number of syllables per clause	high number of syllables per clause
low number of morphological cases	high number of morphological cases
cumulative case exponents	separatist case exponents
stress-timed	syllable-timed
prepositions	postpositions
SVO	SOV
isolating or fusional morphology	agglutinative morphology

Note. Adapted from Fenk-Oczlon & Fenk, 2005

Lupyan and Dale (2010) report an inverse relationship between morphological simplicity (fewer cases, isolating, etc.) and population size. The present paper demonstrates an inverse relationship between phoneme inventory size and morphological simplicity (fewer cases, isolating, etc.). If morphological simplicity correlates with both large population size and large phoneme inventory size, then this might be an explanation for Hay and Bauer's (2007) finding that inventory size and population size are positively correlated.

We propose that a systemic approach to phoneme inventory size will advance our understanding of phoneme inventory size evolution and variation.

References

Atkinson, Q.D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science, 332*, 346–349.

Fenk-Oczlon, G., & Fenk, A. (1999). Cognition, quantitative linguistics, and systemic typology. *Linguistic Typology,* 3, 151-177.

Fenk-Oczlon, G., & Fenk, A. (2005). Crosslinguistic correlations between size of syllables, number of cases, and adposition order. In G. Fenk-Oczlon and C. Winkler (Eds.), *Sprache und Natürlichkeit. Gedenkband für Willi Mayerthaler* (pp.75-86). Tübingen: Narr.

Fenk-Oczlon, G., & Fenk, A. (2008). Complexity trade-offs between the subsystems of language. In M. Miestamo, K. Sinnemäki and F. Karlsson (Eds.), *Language complexity: Typology, contact, change* (pp. 43-65). Amsterdam / Philadelphia: John Benjamins.

Hay, J., & Bauer, L. (2007). Phoneme inventory size and population size. *Language, 83*, 388–400.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure, *PLoS One 5* e8559

Nettle, D. (1995). Segmental inventory size, word length, and communicative efficiency, *Linguistics, 33,* 359-367.

Nettle, D. (1998). Coevolution of phonology and the lexicon in twelve languages of West Africa. *Journal of Quantitative Linguistics*, *5*, 240–245.

Maddieson, I. (2006). Correlating Phonological Complexity: Data and Validation, *Linguistic Typology, 10*, 106–123.

Maddieson, I. (2009). Monosyllables and syllabic complexity. Abstract, Festival of languages, *Monosyllables: From Phonology to Typology*, University of Bremen, Sept. 2009.

Maddieson, I., & Precoda, K. (no date). UCLA Phonological Segment Inventory Database, Electronic database, University of California, Los Angeles. Accessible via interface at http:// web.phonetik.uni-frankfurt.de/upsid.html

Moran, S., McCloy, D., & Wright, R. (eds.) (2014). *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at https://phoible.org

Wichmann., Rama T., & Holman, E. W. (2011). Phonological diversity, word length, and population sizes across languages: The ASJP evidence, *Linguistic*

Typology, 15, 177–197.

TRANSMISSION CHAINS AND INDIVIDUALS PLACE DIFFERENT CONSTRAINTS ON THE EVOLUTION OF CATEGORY SYSTEMS

Vanessa Ferdinand[*1], Charles Kemp[1], and Amy Perfors[1]

[*]Corresponding Author: vanessa.ferdinand@unimelb.edu.au
[1]Melbourne School of Psychological Science, University of Melbourne, Australia

Evolution and learning are both processes that allow organisms to extract and store information about their environment. But when and how do the exploratory dynamics of these processes differ? We present the results of an iterated category learning experiment, where the number and placement of participants' category boundaries are free to evolve over time. We contrast two evolutionary regimes: one where category systems are transmitted over multiple learners and one where they are developed within a single learner, for the same amount of time. We find that there are more constraints on the exploratory process when systems are culturally transmitted among multiple learners. Single learners explore a wider range of category systems and converge on more complex systems, whereas transmission chains explore a more restricted set of systems and nearly always converge on a simple, but easily learnable, one-boundary category system.

297 adults participated in an iterated category learning experiment, where they were trained and tested on a mapping between 2 labels and 10 stimuli. The stimuli had identical shape (a seashell) and varied on a continuous dimension (shade of grey). Participants were divided into two transmission conditions: *cultural* and *individual*, and two frequency conditions: *uniform* and *skewed*. In the *cultural* condition, each participant was trained on a mapping generated directly from the test trials of the previous participant. These participants were organized into 45 linear transmission chains. The first participant in each chain was trained on a random mapping of labels to stimuli. Chains ended when two consecutive participants produced identical mappings in their testing phase (chain length ranged from 3-11 rounds or "generations"). In the *individual* condition, each participant was trained on their own previous test mapping for multiple rounds. Rounds ended when they produced the same mapping twice in a row (or until they completed 8 rounds). In the *uniform* condition, each stimulus was presented an equal number of times (3 times each per training round). In the *skewed* condition, the stimuli were presented with the following frequencies: 10,5,4,3,2,2,1,1,1,1. This condition was counterbalanced between participants so half saw the darker stimuli more frequently and half saw the lighter stimuli more frequently.

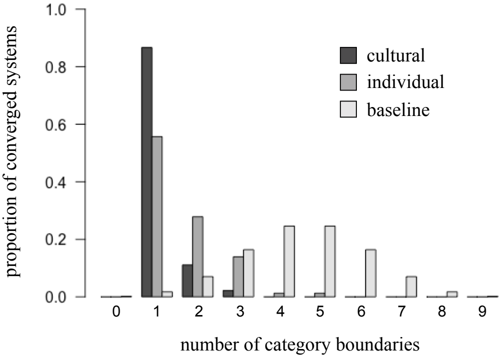


Figure 1.

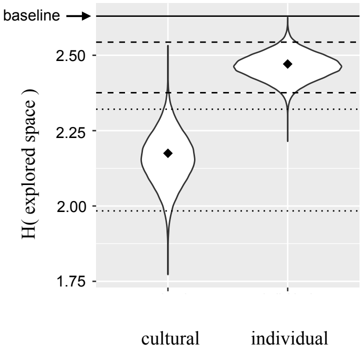


Figure 2.

We found that participants in the cultural condition were significantly more likely to converge on a simple category system with only one boundary, than participants in the individual condition (Figure 1). The baseline in Figure 1 shows the distribution of boundaries that would be produced if participants were randomly clicking on category labels in the testing phase. Figure 2 visualizes the difference in exploratory behavior between conditions using Shannon's 1948 framing of relative entropy. The entropy of the baseline distribution in Figure 1 is 2.6 bits; any distribution with lower entropy has more structure and covers less of the total space of possible category systems. The entropy of the systems that participants explored is significantly lower, at 2.47 bits in the individual condition and 2.17 bits in the cultural condition.[1] The dashed lines show the 95% confidence intervals[2] around the entropy estimate: their lack of overlap indicates that participants in the cultural condition explored a significantly smaller region of all possible category systems than participants in the individual condition.

As for the frequency manipulation, we predicted that it would constrain evolutionary search by affecting the location of category boundaries, but these analyses were confounded by a significant difference in the counterbalance condition: participants were more likely to place boundaries in the lighter region of the greyscale, regardless of skew condition. Future research will focus on unpacking the observed limits on evolutionary search into identifiable constraints that differed between conditions (e.g. the differing demands on memory and attention).

---

[1] Exploratory breadth was obtained by estimating the Shannon entropy of the distribution of category systems obtained in all testing phases, up to an evolutionary depth of 8 rounds/generations. (Systems in the cultural condition were allowed to evolve for more than 8 generations, so this cut-off aids comparability between conditions). All entropy calculations were corrected with minimax estimation (Hausser & Strimmer, 2014).

[2] Computed with the bootstrap percentile method (Efron, 1979).

References

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The annals of Statistics*, 1–26.

Hausser, J., & Strimmer, K. (2014). *entropy: Estimation of entropy, mutual information and related quantities.* (R package version 1.2.1)

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, *27*(3), 379–423.

LIMITATIONS IN AVIAN PERCEPTION OF VOCAL SEQUENCES: IMPLICATIONS FOR LANGUAGE EVOLUTION

ADAM R. FISHBEIN*[1], NORA H. PRIOR[1], GREGORY F. BALL[1] and ROBERT J. DOOLING[1]

*Corresponding Author: afishbei@umd.edu

[1]Department of Psychology, Neuroscience and Cognitive Science Program, University of Maryland, College Park, USA

Birdsong and human speech have strong parallels: both consist of learned acoustic elements produced in sequences (Bolhuis, Okanoya, & Scharff, 2010; Doupe & Kuhl, 1999). In speech, sequences are fundamentally important in the perception of phonological and syntactic patterns. In birdsong, the production of sequences has been well-described, but the importance of these sequences in perception is uncertain or equivocal for most species. Here, we summarize the results of several studies which show limitations in avian perception of sequences and a greater importance of the acoustic details of individual elements. We also highlight greater sequence sensitivity in the budgerigar (Melopsittacus undulatus), a psittacine, compared to the zebra finch (Taeniopygia guttata), a songbird, illustrating how a careful species comparison might be useful for studying human language evolution.

In our studies, birds were trained through operant conditioning to perform a psychophysical discrimination task. The bird pecks a key (the observation key) while listening to a repeating sound (the background stimulus) in order to elicit either a different sound (a target stimulus) or the same sound (a sham trial). If the bird pecks another key (the report key) when the target sound is presented, then it is rewarded with millet from a food hopper.

In study 1, three songbird species (zebra finch, canary, and Bengalese finch) were tested along with a psittacine (budgerigar) (Lawson, Fishbein, Prior, Dooling, & Ball, 2018). In this study, we used a zebra finch song motif (consisting of 3-8 elements or "syllables") as the background, while the targets involved either song elements temporally reversed or song elements shuffled out of their natural sequence. The results showed that all the species tested were very sensitive to reversals of individual elements, but that budgerigars were much better than the songbirds at hearing changes to sequence. The zebra finches' relative lack of sensitivity to the sequence of elements in conspecific motifs is surprising given that these birds learn to produce their song elements in a fixed sequence early in life and maintain the sequence throughout life.

Budgerigar song, called "warble," consists of a rambling, jazz-like structure. Previous work has shown that the birds are sensitive to violations in the sequential order of natural song (Tu & Dooling, 2012), but it is not clear what sort of rules they can use: (1) rules about the transitions between particular sounds or (2) rules about the abstract relationships among elements. In study 2, budgerigars were tested on their ability to discriminate changes to a sequential pattern of conspecific song elements, AAB, i.e. same-same-different. In experiment 1, budgerigars, unlike humans, primarily used transition rules to detect violations of the AAB pattern, rather than abstract structure. In experiment 2, the stimuli were presented so that transition rules were no longer reliable cues, and the birds showed the ability to use abstract structure but only of the first two elements.

During song bouts, zebra finches sing many renditions of their motifs, which have the same sequential structure but differ in fine-grained acoustic details. This dimension of their song has been largely overlooked by researchers, as the sequential structure of their song is more striking to casual human listeners and was assumed to be a salient song feature for zebra finches. In study 3, zebra finches were tested on their ability to discriminate the subtle acoustic differences that occur among renditions of the same song syllables by the same individual. We show that zebra finches are very sensitive to subtle acoustic variation in renditions of natural song motifs which is in striking contrast to their insensitivity to sequence changes in study 1.

Together these experiments highlight the limitations birds have in perceiving sequence information, which is in striking contrast to human language perception. Birdsong may be an architecturally different communication system than human language – where information in birdsong is primarily conveyed in the acoustic details of vocal elements with the sequential patterns of those elements having at best secondary importance. Yet, the species differences between zebra finches and budgerigars in sequence perception observed here and in other work (Spierings & Ten Cate, 2016) offer an important opportunity for comparative neurobiological studies. Explaining how budgerigar brains are better able to process acoustic sequences than songbird brains could provide clues as to why human brains are better adapted for sequence processing than those of other primates.

Acknowledgements

AF is supported by funding from NIDCD F31-DC017884, T32-DC000046, and the NSF under grant No.1449815.

References

Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews Neuroscience*, *11*, 747–759. https://doi.org/10.1038/nrn2931

Doupe, A. J., & Kuhl, P. K. (1999). Bird Song and Human Speech: Common Themes and Mechanisms. *Annu. Rev. Neurosci.*, *22*, 567–631.

Lawson, S. L., Fishbein, A. R., Prior, N. H., Dooling, R. J., & Ball, G. F. (2018). Relative salience of syllable structure and syllable order in zebra finch song. *Animal Cognition*, *21*(4), 467–480. https://doi.org/10.1007/s10071-018-1182-2

Spierings, M. J., & Ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proc Natl Acad Sci U S A*, *113*(27), 3977–3984. https://doi.org/10.1073/pnas.1600483113

Tu, H.-W., & Dooling, R. J. (2012). Perception of warble song in budgerigars (Melopsittacus undulatus): Evidence for special processing. *Animal Cognition*, *15*(6), 1151–1159. https://doi.org/10.1007/s10071-012-0539-1

AN EMERGENT LANGUAGE BECOMES SMALLER AS IT EVOLVES: NEW EVIDENCE FROM MOTION TRACKING IN NICARAGUAN SIGN LANGUAGE

MOLLY FLAHERTY*[1], ASHA SATO[2], and SIMON KIRBY[2]

*Corresponding Author: mflaher1@swarthmore.edu
[1]Department of Psychology, Swarthmore College, Swarthmore PA, USA
[2]Centre for Language Evolution, University of Edinburgh, Edinburgh, UK

With the founding of a new school in Managua approximately forty years ago, Deaf Nicaraguans came together in greater numbers than before. Though teaching was in Spanish, Deaf students soon began to communicate manually. This was the beginning of a new language: Nicaraguan Sign Language (NSL). Each year children enter the school and learn the language from their older peers, eventually becoming Deaf adults who use NSL for daily communication. Over successive generations, the language grows and changes. By comparing signers of different generations, we can document these changes. Here we investigate this structural change with motion tracking technology. Using the Kinect Motion Sensor (similar to Namboodiripad et al., 2016), we quantify a change that has often been hypothesized to occur in the development of a new sign language: reduction in size of the signing space.

Decrease in the size of a language's signing space, the area in front of the body in which signers produce signs, has been cited as evidence of maturation of a language and development away from gestural roots (e.g., Nyst, 2007). This phenomenon has previously been mentioned in NSL (Kegl et al., 1989), but in the twenty years since that first mention it has not been rigorously quantified.

Seventeen Deaf Nicaraguan Signers participated, all of whom began signing in childhood. Participants were drawn from a wide age range, having entered the Deaf Community from soon after the founding of the school (1974) to nearly thirty years later (2003). During an elicitation task, we tracked the position of signers' wrists using the Kinect, which returns inferred XYZ positions of 21 joints at a target frame rate of 30 fps (Schotton et. al., 2013). Before analysis, skeleton data for each participant was filtered using median filtering to reduce noise, and skeletons were scaled to have the same upper-arm length to minimize effect of differing body proportions.

We measured signing space using the Euclidean distance between the tracked position between the shoulders and the tracked position of each wrist. This distance was computed at each frame and averaged over the session to obtain the mean distance of the wrists to the base of the shoulders for each signer. A simple linear model predicting size of signing space as a function of year of entry finds a significant effect (*F(1,15)=6.15, p=0.025*): overall signing space decreases with later year of entry. As this measure characterizes all of the sign and non-sign movements produced by a participant over the entire session, it is potentially subject to substantial noise. To address this concern, we isolated individual utterances and repeated the same analysis on those utterances alone. Here we found the same pattern (Figure 1), signers who entered the NSL community later employ a smaller signing space than do older signers (*F(1,14)=7.43, p=0.016).*

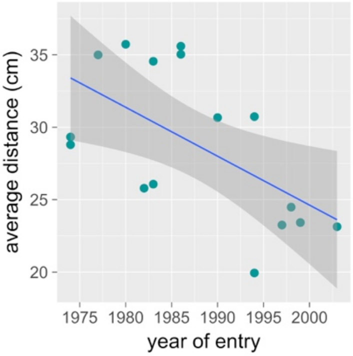


Figure 1. Normalized average distance from the base of the shoulder to the wrists for each signer during verb production. As the language matures, the size of the signing space decreases.

These results show that, even correcting for body size, younger signers who have learned NSL after it has been evolving longer employ a smaller signing space than older signers, reflecting a change in the language itself: signing space in NSL is decreasing in size as it is passed down to subsequent generations. Using motion tracking data, we have documented that as signers learn, use, and transmit NSL, the language begins to more closely resemble older world sign languages.

Acknowledgements

Supported by a British Academy NIF and Alumni NIF (MF), Economic and Social Research Council, grant number ES/J500136/1 (AS), and ESRC grant ES/R011869/1 (SK).

References

Kegl, J., & Iwata, G. (1989). Lenguaje de Signos Nicaragüense: A pidgin sheds light on the “creole?” ASL. In *Proceedings of the Fourth Annual Meeting of the Pacific Linguistics Conference* (pp. 266-294). Eugene: University of Oregon, Department of Linguistics.

Namboodiripad, S., Lenzen, D., Lepic, R., & Verhoef, T. (2016). Measuring conventionalization in the manual modality. *Journal of Language Evolution*, *1*(2), 109-118.

Nyst, V. A. S. (2007). *A descriptive analysis of Adamorobe sign language (Ghana)* Utrecht, The Netherlands: LOT.

NATURAL POPULATION GROWTH CAN CAUSE LANGUAGE SIMPLIFICATION

Stella Frank*[1] and Kenny Smith[1]

*Corresponding Author: stella.frank@ed.ac.uk
[1]Centre for Language Evolution, University of Edinburgh, UK

Linguistic complexity is known to be negatively correlated with population size. We hypothesise that language simplification could occur during natural population growth as a result of increased numbers of learners in the population. Simulation results confirm this hypothesis: younger age distributions result in lower complexity, independently of population size, and growing populations show a drop in complexity that matches the increase in young learners.

1. Introduction

Larger populations tend to speak simpler languages, as measured by the diversity of morphological and grammatical structures (Lupyan & Dale, 2010). This population size effect has been argued to be a result of population growth via immigration, with adult learners having a simplifying effect on the language (Wray & Grace, 2007; McWhorter, 2007; Trudgill, 2011, see also Bentz & Winter (2013); Bentz et al. (2015) for empirical support). However, in a large-scale analysis, Koplenig (2019) does not find an effect of language "vehicularity" (whether a language is learned by adults) on complexity, but does find an effect of population size. This recent result requires reexamining the proposed mechanism behind the link between population size and linguistic complexity.

An alternate route to a larger population that does not involve immigration (with the attendant adult language learners), is so-called 'natural' population growth as the result of birth rates exceeding death rates. More specifically, if within the population speaking a given language, there are more new speakers (due either to increasing birth rates or decreasing infant and early-age mortality rates), while adult mortality remains constant, the total number of speakers of that language will increase. The demographic consequences of natural population growth is an age distribution that is skewed towards the young, a phenomenon found in contemporary growing populations but also in small-scale societies in the past. Data from modern hunter-gatherer and other small-scale societies indicate that the main source of demographic variation between these groups is the rate of infant and early childhood mortality while adult rates of mortality are relatively similar (Gurven & Kaplan, 2007; Pennington, 1996). Increases in population size, whether enabled by climate or other factors (Tallavaara & Sepp, 2011; Bettinger, 2016), would thus be most likely driven by more children surviving to adulthood, rather than by adults living longer. During the Neolithic transition to agriculture, population demographics changed dramatically in what is called the Neolithic Demographic Transition (Bocquet-Appel, 2011; Shennan, 2001): birth rates increased sharply, and were only later balanced out by increased mortality rates. This led to youth-heavy populations: "At the peak of the NDT, there were children everywhere and the average age of the population was about 18 years old" (Bocquet-Appel, 2011).

What effect could this youth-heavy demographic distribution have on language? Younger speakers still in the process of learning the language are generally simply less accurate speakers of the language. Language changes are also often led by younger speakers, amplified by networks of age-peers learning from and reinforcing one another (Labov, 2007; Cournane, 2017; Sankoff, 2018). A population with more younger speakers, such as a naturally growing population, might thus be subject to more linguistic change, or at least more attempted changes; this could also lead to fewer fully accepted variants, if new idiosyncratic variants displace older, more wide-spread variants in learner's inputs.

In this paper, we use an agent-based population model to demonstrate that the above holds: increasing the proportion of younger speaker agents leads to a drop in the number of linguistic variants shared by every agent in the population, i.e., the simplest measure of linguistic complexity. These shared variants can be thought of as the set of forms that are a required part of a speaker's linguistic inventory. The agents within the model are formulated to be comparable with previous work (Reali et al., 2014, 2018; Spike, 2017) but we introduce more realistic population dynamics, including population turnover.

In our simulations, we first disentangle population size and population demographics by simulating the development of languages in populations with different demographic distributions, e.g. populations in which there are always more younger speakers than older speakers, or in which ages are more evenly distributed, but without population growth. We find that stationary populations with more younger speakers converge to languages with lower complexity, as measured by the number of variants shared throughout the population. Age demographics and size interact, with younger smaller populations having similar complexity levels as older larger populations.

In actually growing populations, in which more agents are added than removed, we then find that language complexity drops as the population increases. The drop in complexity is in proportion to the rate of population growth, i.e., the extent to which the demographic distribution shifts from being evenly distributed to more youth-heavy. This further supports our hypothesis that increasing numbers of younger speakers, due to natural population growth, is an alternate mechanism to language simplification.

2. Model details

2.1. *Agents and their language*

We follow Reali et al. (2014, 2018); Spike (2017) in using a Dirichlet Process or (equivalently) Hoppe Urn model of agent language learning. Agents learn a set of items (corresponding to lexical items or grammatical variants) by interacting with other agents. When they hear an item, it gets added to their lexicon. Agents speak by drawing from the lexicon, where the probability of producing an item is proportional to the number of times (tokens) that item (type) has been stored in the lexicon: $p(x) = \text{Count}(x)/(M+\alpha)$, where M is the number of tokens seen so far and α is an 'innovation' hyperparameter. Namely, agents also have the possibility of inventing a new item, with probability inversely proportional to the number of tokens already in their lexicon: $p_{new} = \alpha/(M + \alpha)$. We set $\alpha = 1$ throughout. Innovated items are always unique (two agents can't separately innovate the same item). Agents also update the counts in their lexicon with their own productions, in order to ensure that their own new innovations are part of their stored lexicon. At the very beginning of the simulation, there are no seeded items: the agents start out with empty lexicons, the same as when agents enter the population later on.

The agents' lexicons may be bounded by a memory limit, in which case only the last m tokens (heard or spoken) are retained in the lexicon and the oldest tokens are deleted. Rare types will thus disappear if they are not used. Note that the memory limit also indirectly sets the lower bound on the probability of generating new items, since M is upper bounded by the memory limit.

Because of their lack of experience, and subsequent smaller M, younger agents will be more influenced more by every interaction they have, and are more likely to adopt new variants, as well as innovate new variants themselves. Older agents on the other hand tend to be more conservative than younger agents, in the sense that when they encounter a new variant, they will add it to their lexicon, but will be unlikely to start producing it themselves. A memory limit will limit the convergence of older agents, and consequently reduce the difference between older and younger agents.

2.2. *Agent interactions*

Agents interact in dialogues in which each agent speaks and then listens to the other agent for 10 turns.[1] Agents participate in 10 dialogues per 'epoch', i.e. between turnover periods, described in the next section.

The population is organised as a fully-connected graph, meaning the probability of any two agents being paired up is equal. As a consequence, young agents in

---

[1] We found longer dialogues, while not common in previous work, to be important for getting new variants to spread: they need to build up enough probability in the context of first use, in a kind of 'conceptual pact', for the agents to reuse them in subsequent dialogues with other agents.

youth-heavy populations will thus be more likely to speak among themselves and will, as a result, learn language as much (or more) from their age peers than from their elders.

Agents are considered to be adults after an initial learning phase consisting of two epochs, corresponding to 400 updates in 20 dialogues. During this initial learning phase, child agents do not speak, so other agents do not update, analogous to adults ignoring child babbling.

The complexity of a language in a population at a given point in time is measured as the number of variants shared by all adult agents in the population, i.e. the size of the common language. This is the same criterion as used by Spike (2017) and we use it to avoid setting arbitrary thresholds. We have verified that the pattern of results is the same with less stringent criteria, e.g. measuring complexity as the number of variants shared by 50% of the population.

2.3. *Population Turnover*

Figure 1. Population pyramids created by removing agents at different turnover rates (columns) using the Gompertz function with two different β settings (rows). Increased turnover leads to younger populations, with a lower mean age.

We implement gradual turnover within our population by selecting a number of agents at each epoch to replace. The turnover parameter p regulates how many agents are removed at each epoch ($p \times N$, the number of agents in the population); in a stable (not growing) population, the same number of agents are replaced with new agents. The more agents are replaced, the more the resulting demographic distributions are weighted towards younger agents. Figure 1 shows these age distributions in the form of 'population pyramids': higher levels of turnover lead to larger numbers of younger agents and a lower average age of the population.

In a realistic population, older agents are more likely to be removed than younger agents. To achieve this dynamic, we use the Gompertz function, which resembles an asymmetric sigmoid and was developed to characterise human life expectancy at higher ages (Gompertz, 1825, see also Baxter & Croft (2016) for use in a model of language change). We follow the parameterisation given in Missov et al. (2015), where the age-dependent probability of picking an agent for removal is $p(age) = \beta \exp \beta(age - M)$, where β is a rate parameter and M is a parameter corresponding to modal age at death (set throughout to 100, although in populations with high turnover, effective age at death will be much younger). Changing the rate parameter β affects the likelihood of picking only the oldest agents: with a smaller setting ($\beta = 0.01$), younger agents are also sometimes removed, leaving some older agents to remain in the population longer.

Importantly, we do not select the agents to remove uniformly at random (cf. Spike, 2017; Dale & Lupyan, 2012; Reali et al., 2018), since this leads to an unrealistic age distribution with mostly young agents but a small number of very long-lived agents. (More formally, age is exponentially distributed as a result of the Poisson point process governing removal.)

2.4. *Related Work*

Agent-based models of language evolution have replicated the effect of population size on language complexity (Dale & Lupyan, 2012; Reali et al., 2014, 2018; Spike, 2017). However, these models either do not include population turnover or do turnover at random, which we argue leads to unrealistic demographic distributions. Similarly, models of sociolinguistic variation and the dynamics of language change (Baxter & Croft, 2016; Stanford & Kenny, 2013; Kauhanen, 2016) have added more realism to network structure and population turnover, but have not investigated the effect of demographic distributions.

In some iterated learning models, populations are modelled as a series of non-overlapping discrete generations, with (vertical) transmission exclusively from older to younger generations (Griffiths & Kalish, 2007; Griffiths & Reali, 2011; Kirby et al., 2014). In contrast, in our population agents enter the population continuously, and learn from interactions. Learning is also symmetrical: within a given interaction, both agents are updating their language, not only the younger agent (though for the reasons spelled out above, the older agent will be influenced less by the younger agent than vice-versa). As a consequence agents can influence other agents that are the same age (horizontal transmission) or older. Another important difference is that, in contrast to ‘chain’ models in which an agent learns only from a single other (older) cultural parent agent, the agents in our population learn from multiple speakers (see also Niyogi & Berwick, 2009; Smith, 2009; Burkkett & Griffiths, 2009).

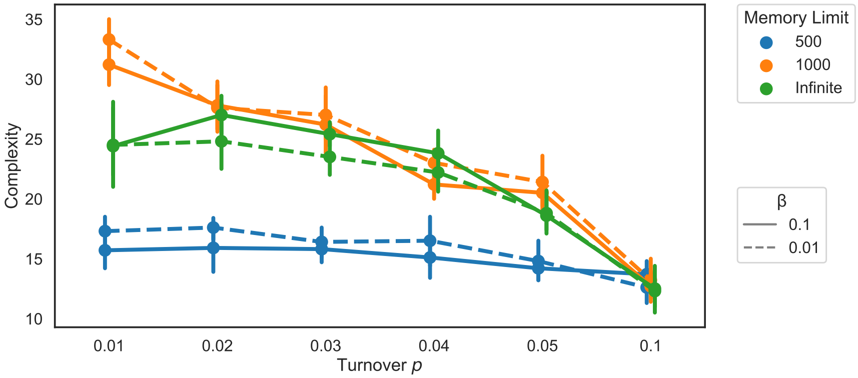


Figure 2. Final complexities of populations with different turnover rates (x- axis), β parameter settings (line styles), and memory limits (colors). Populations become younger going from left to right, in line with a decrease in complexity. Averages shown are over 10 runs.

3. Simulation Results

3.1. *Stable populations with different age distributions*

In our first experiment, we keep the population size stable (at $N = 100$) and vary only the population age distribution by manipulating the rate of turnover. While our broader argument is about population growth driven by an increase in the proportion of young agents, a stable non-growing population allows us to separate the effect of age demographics from population growth. In this setting, any differences in linguistic complexity between populations of the same size are due to the relative numbers of young and older agents.

We run each population simulation for 10000 epochs and report the complexity of the language at the final epoch. The final complexities for the populations in Figure 1 with three different memory limits are show in Figure 2. Higher turnover rates, resulting in younger populations, lead to languages with fewer shared variants. Conversely, older populations have languages with more shared variants, as long as the agents have sufficient memory capacity. With a low memory limit all agents act ‘young’ and complexity is reduced across all turnover rates.

Linguistic complexity thus requires a sufficient proportion of experienced older agents with in the population. The exact shape of the age distribution matters less, since the value of β, which mainly effects the number of older agents, doesn’t have an effect. When the numbers of old and young are not balanced, the young learn from peer interactions rather than from older agents. These populations are characterised by more innovative variants which do not spread through the whole population and thus do not contribute to complexity as measured here.

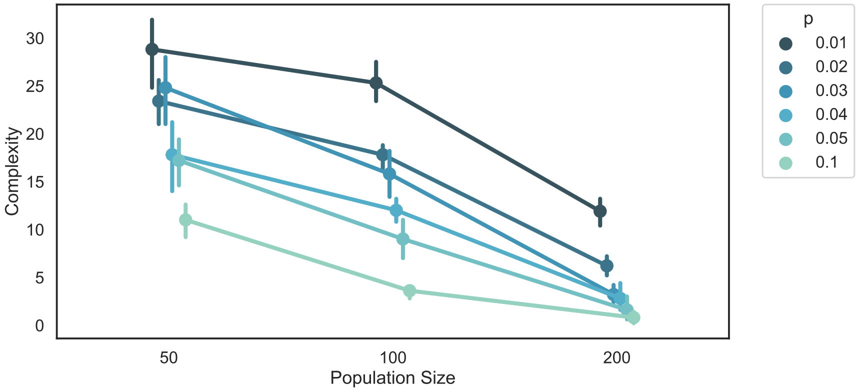


Figure 3. Final complexities of populations of different sizes (x-axis) and demographics (lines). Memory limit $= 1000$, Gompertz $\beta = 0.1$.

3.2. *Population size and age demographics*

The effect that older populations (lower turnover) have higher complexity holds across population sizes (shown in Figure 3, which also shows a replication of the finding that smaller populations have higher complexity across demographics). Intriguingly, smaller youth-heavy populations (i.e. higher p, corresponding to the lower lines on the graph) have similar complexity to larger populations with more evenly distributed age demographic: the former have the demographic signature of growing populations, while the latter are more stable, suggesting a constant level of complexity from a small but growing population to the more stable larger population that is the result of growth.

3.3. *Population growth*

Finally, we check that actual population growth affects the language of the population. In these simulations, we first run a fixed-sized population for long enough for the language to stabilise, at a turnover rate ($p = 0.01$) that results in a population with an even age distribution. During the growth phase, the rate of adding agents is larger than the rate of removing agents by the growth factor g. Depending on g, growth happens faster or slower; we stop growth after the population size has quadrupled.

Population growth leads to an immediate decrease in complexity (Figure 4), with higher rates of growth leading to larger decreases. When growth stops and the population stabilises, complexity increases again, but crucially, at a lower level than before population growth. Natural population growth can thus capture the decrease in complexity linked to larger populations.

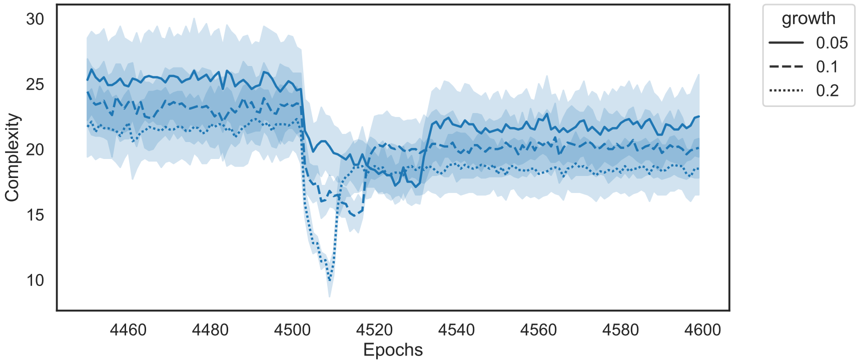


Figure 4. Complexity of a population growing at different rates (g). Growth starts at Epoch 4500 and stops when the population has quadrupled, i.e., $N \geq 400$. Initial $p = 0.01$, no memory limit (infinite), Gompertz $\beta = 0.1$, 10 runs of each setting. Note the initial differences (before Epoch 4500) between g values are due to random fluctuations and are not meaningful.

4. Conclusion

Larger populations can be the result of either migration or natural population growth. Growing populations are characterised by an increased proportion of younger members as compared to stable populations. In our model, populations with these characteristics achieve lower complexity in their linguistic system than populations with the even age distribution associated with stable populations. This offers an alternative route to explaining the link between population size and language complexity that does not involve non-native speakers.

Our results are a consequence of the different learning environments of agents in youth-heavy vs. stable, older, populations. In stable populations, the youngest agents will interact mostly with older agents from whom they learn the language shared by the rest of the population. In contrast, in youth-heavy populations, young agents interact more with their age peers, who have similarly not been exposed to the full language, and thus complexity is lost. The fully-connected network assumed by our model is unrealistic for human populations which have more differentiated social networks. However, in small-scale networks, fully connected networks have similar characteristics to more realistic small-world networks (Spike, 2017). In exploratory modelling experiments, we found that early learner network connectivity (e.g. learning from a principal caretaker) could not mitigate the demographic effect presented in this paper. Secondly, the argument here in a relative one: are children in youth-heavy populations interacting more with, and learning more from, their peers than children in older populations? The answer to this is likely to be yes, though will depend on culturally-specific child-rearing practices and institutions.

References

Baxter, G., & Croft, W. (2016). Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change*, *28*(02), 129–173.

Bentz, C., Verkerk, A., Kiela, D., Hill, F., & Buttery, P. (2015). Adaptive communication: Languages with more non-native speakers tend to have fewer word forms. *PLOS ONE*, *10*(6), e0128254.

Bentz, C., & Winter, B. (2013). Languages with more second language learners tend to lose nominal case. *Language Dynamics and Change*, *3*, 1–27.

Bettinger, R. L. (2016). Prehistoric hunter–gatherer population growth rates rival those of agriculturalists. *Proceedings of the National Academy of Sciences*, *113*(4), 812–814.

Bocquet-Appel, J.-P. (2011). When the world's population took off: The springboard of the Neolithic demographic transition. *Science*, *333*(6042), 560–561.

Burkkett, D., & Griffiths, T. L. (2009). Iterated learning of multiple languages from multiple teachers. In *Proceedings of Evolang.*

Cournane, A. (2017). In defense of the child innovator. In Éric Mathieu & R. Trueswell (Eds.), *Micro-change and macro-change in diachronic syntax* (chap. In defence of the child innovator). Oxford University Press.

Dale, R., & Lupyan, G. (2012). Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems*, *15*(03n04), 1150017.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, *115*(0), 513–583.

Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, *31*(3), 441–480.

Griffiths, T. L., & Reali, F. (2011). Modelling minds as well as populations. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1713), 1773–1776.

Gurven, M., & Kaplan, H. (2007). Longevity among hunter-gatherers: A cross-cultural examination. *Population and Development Review*, *33*(2), 321–365.

Kauhanen, H. (2016). Neutral change. *Journal of Linguistics*, *53*(02), 327–358.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

Koplenig, A. (2019). Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *Royal Society Open Science*, *6*(2), 181274.

Labov, W. (2007). Transmission and diffusion. *Language*, *83*(2), 344–387.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE*, *5*(1), e8559.

McWhorter, J. (2007). *Language interrupted: Signs of non-native acquisition in standard language grammars*. Oxford University Press.

Missov, T. I., Lenart, A., Nemeth, L., Canudas-Romo, V., & Vaupel, J. W. (2015). The Gompertz force of mortality in terms of the modal age at death. *Demographic Research*, *32*, 1031–1048.

Niyogi, P., & Berwick, R. C. (2009). The proper treatment of language acquisition and change in a population setting. *Proceedings of the National Academy of Sciences*, *106*(25), 10124–10129.

Pennington, R. L. (1996). Causes of early human population growth. *American journal of physical anthropology*, *99*, 259–274.

Reali, F., Chater, N., & Christiansen, M. H. (2014). The paradox of linguistic complexity and community size. In *The evolution of language: Proceeedings of the 10th international conference.*

Reali, F., Chater, N., & Christiansen, M. H. (2018). Simpler grammar, larger vocabulary: How population size affects language. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1871), 20172586.

Sankoff, G. (2018). Language change across the lifespan. *Annual Review of Linguistics*, *4*(1), 297–316.

Shennan, S. (2001). Demography and cultural innovation: a model and its implications for the emergence of modern human culture. *Cambridge Archaeological Journal*, *11*(01).

Smith, K. (2009). Iterated learning in populations of Bayesian agents. In *Proceedings of CogSci.*

Spike, M. (2017). Population size, learning, and innovation determine linguistic complexity. In *Proceedings of CogSci.*

Stanford, J. N., & Kenny, L. A. (2013). Revisiting transmission and diffusion: An agent-based model of vowel chain shifts across large communities. *Language Variation and Change*, *25*(02), 119–153.

Tallavaara, M., & Sepp, H. (2011). Did the mid-holocene environmental changes cause the boom and bust of hunter-gatherer population size in eastern fennoscandia? *The Holocene*, *22*(2), 215–225.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford University Press.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543–578.

MULTIMODAL AND MULTIPLEX COMMUNICATION IN WILD AND CAPTIVE ORANG-UTANS (*PONGO ABELII/PYGMAEUS*)

MARLEN FRÖHLICH[*1] and CAREL P. VAN SCHAIK[1]

[*]Corresponding Author: marlen.froehlich@uzh.ch
[1]Department of Anthropology, University of Zurich, Zurich, Switzerland

The ecological niche of language is face-to-face interaction, hence human communication is inherently multimodal (Holler & Levinson, 2019). Non-human primates and many other species also have a natural predisposition for multimodal social interactions. With regard to the function(s) of multimodality in the animal kingdom, hypotheses have invoked redundancy, refinement, or complementarity. However, very few studies addressed this issue in great apes, our closest living relatives. Similar to humans, great apes display striking behavioural plasticity in the communicative domain, and we might gain critical insights about learning effects by studying the flexible combination of modalities (i.e. sensory channels) and articulators (i.e. communication organs) in different socio-ecological environments.

Compared to the African great apes, orang-utans' (*Pongo* spp.) non-vocal communication has to date been little studied in natural environments (but see Fröhlich et al., 2019; Knox et al., 2019; MacKinnon, 1974; Rijksen, 1978). However, studies over the past decades have shown that populations of the two orang-utan species on Borneo (*Pongo pygmaeus*) and Sumatra (*Pongo abelii*) differ considerably in sociability and behavioural variants such as call types and tool-use techniques, which has been interpreted as evidence for culture (van Schaik et al., 2003; Wich et al., 2012). In contrast, in captive settings orang-utans are often exposed to a social and semi-terrestrial lifestyle similar to that of wild chimpanzees and bonobos. Specifically, they encounter more opportunities for social interactions with conspecifics other than the mother, resulting in a large proportion of well-established interaction dyads and a lower risk of misunderstanding, as well as better visibility and a more terrestrial lifestyle. Orang-utans in captivity show substantially higher levels of social interactions than observed in the wild, including social play and grooming sessions among adults and socio-sexual behaviours among immature individuals (Poole, 1987). These population and setting (wild-captive) contrasts make orang-utans eminently suitable to test hypotheses about the function of multimodality in apes and the effects of learning on its incidence.

In the present study, we examined the communicative behaviour of Bornean and Sumatran orang-utans living in the wild (*Tuanan*, *Suaq*) and in European zoos (*Apenheul, Cologne, Munster, Munich, Zurich*), to explicitly test to what extent socio-ecology—particularly, visibility and availability of social partners—influenced communicative behaviour. During ca. 1,600 hours of focal observations, we video-recorded more than 6,300 communicative interactions, resulting in the coding of around 10,000 communicative acts. Here, we focused specifically on the use of multimodal (i.e. involving multiple sensory modalities) as well as multiplex (i.e. involving multiple articulators, such as mouth and limbs) communication. Our 2x2 comparative design allowed us to disentangle effects of species and research setting while controlling for age and sex as well as contextual and recipient-related factors. Our first analyses of this extensive dataset revealed that purely visual (non-contact) communicative acts were significantly less, but tactile and multimodal acts more common in *P. pygmaeus* in the least sociable wild population (Tuanan) compared to the other species-setting combinations. Moreover, both multimodal and multiplex communication was more pervasive in wild than captive settings. Irrespective of these effects of setting and species, we found differences in use between age classes and contexts.

These preliminary findings indicate that the main function of multi-component communication in orang-utans is to facilitate detection and disambiguation of a specific message rather than to carry multiple messages, supporting a redundancy function. As previously shown for wild chimpanzees, predictable outcomes and high visibility may foster the use of unimodal signals when unimpeded face-to-face communication is possible. Communicative efforts are thus reduced like in humans, when interactions between partners occur repeatedly (Clark & Brennan, 1991), or when partners share an extensive interactional history (Brown & Levinson, 1987). Moreover, this also suggests that humans evolved a novel function of multimodality to support more complex messages.

References

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge, MA, US: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (Vol. 13, pp. 127-149). Washington, DC, US: American Psychological Association.

Fröhlich, M., Lee, K., Mitra Setia, T., Schuppli, C., & van Schaik, C. P. (2019). The loud scratch: a newly identified gesture of Sumatran orangutan mothers in the wild. *Biology Letters, 15*(7), 20190209.

Holler, J., & Levinson, S. C. (2019). Multimodal language processing in human communication. *Trends in Cognitive Sciences, 23*(8), 639-652.

Knox, A., Markx, J., How, E., Azis, A., Hobaiter, C., van Veen, F. J. F., & Morrogh-Bernard, H. (2019). Gesture use in communication between mothers and offspring in wild orang-utans (*Pongo pygmaeus wurmbii*) from the Sabangau Peat-Swamp Forest, Borneo. *International Journal of Primatology, 40*(3), 393-416.

MacKinnon, J. (1974). The behaviour and ecology of wild orang-utans (Pongo pygmaeus). *Animal Behaviour, 22*(1), 3-74.

Rijksen, H. D. (1978). *A fieldstudy on Sumatran orang utans (Pongo pygmaeus abelii, Lesson 1827): Ecology, behaviour and conservation*. Netherlands: H. Veenman

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., . . . Merrill, M. (2003). Orangutan cultures and the evolution of material culture. *Science, 299*, 102-105.

Wich, S. A., Krützen, M., Lameira, A. R., Nater, A., Arora, N., Bastian, M. L., . . . Pamungkas, J. (2012). Call cultures in orang-utans? *PLoS ONE, 7*(5), e36180.

COGNITIVE CONSTRAINTS BUILT INTO FORMAL GRAMMARS: IMPLICATIONS FOR LANGUAGE EVOLUTION

CARLOS GÓMEZ-RODRÍGUEZ[1], MORTEN H. CHRISTIANSEN[2,3], and RAMON FERRER-I-CANCHO[*4]

[*]Corresponding Author: rferrericancho@cs.upc.edu
[1]FASTPARSE Lab, LyS Research Group, Departamento de Computación, Universidade da Coruña, CITIC, Campus de Elviña s/n, 15071, A Coruña, Spain
[2]Department of Psychology, Cornell University, Ithaca, NY 14853, USA
[3]Interacting Minds Centre and School of Communication and Culture, Nobelparken, Aarhus University, 8000 Aarhus C, Denmark
[4]Departament de Ciències de la Computació, Universitat Politècnica de Catalunya (UPC), Jordi Girona 1-2, 08034, Barcelona, Catalonia

An often celebrated aspect of human language is its capacity to produce an unbounded number of different sentences (Chomsky, 1965; Miller, 2000). For many centuries, the goal of linguistics has been to capture this capacity by a formal description—a grammar—consisting of a systematic set of rules and/or principles that determine which sentences are part of a given language and which are not (Bod, 2013). Over the years, these formal grammars have taken many forms but common to them all is the assumption that they capture the idealized linguistic competence of a native speaker/hearer, independent of any memory limitations or other non-linguistic cognitive constraints (Chomsky, 1965; Miller, 2000). These abstract formal descriptions have come to play a foundational role in the language sciences (Hauser, Chomsky, & Fitch, 2002; Pinker, 2003). Despite evidence that processing difficulty underpins the unacceptability of certain sentences (Morrill, 2010; Hawkins, 2004), the cognitive independence assumption that is a defining feature of linguistic competence has not been examined in a systematic way using the tools of formal grammar. It is therefore unclear whether these supposedly idealized descriptions of language are free of non-linguistic cognitive constraints, such as memory limitations.

If the cognitive independence assumption should turn out not to hold, then it would have wide-spread theoretical and practical implications for our understanding of human linguistic productivity. It would require a reappraisal of key parts of linguistic theory that hitherto have posed formidable challenges for explanations of language processing, acquisition and evolution (Gold, 1967; Hauser et al., 2002; Pinker, 2003)—pointing to new ways of thinking about language that may simplify the problem space considerably (Christiansen & Chater, 2008; Gómez-Rodríguez & Ferrer-i-Cancho, 2017). Here, we therefore evaluate the cognitive independence assumption using a state-of-the-art grammatical framework, dependency grammar (Nivre, 2005), to search for possible hidden memory constraints in these formal, idealized descriptions of natural language. To delimit the set of possible grammatical descriptions, various classes or sets of syntactic dependency structures have been proposed. These classes can be seen as filters on the set of all the possible syntactic structures. Here, we consider projective structures and various classes of mildly non-projective structures (Gómez-Rodríguez, 2016).

We validate the assumption of independence between grammatical constraints and cognitive limitations in these classes of grammar using the distance between syntactically related words in a dependency tree as a proxy for memory constraints (Liu, Xu, & Liang, 2017). For a given sentence length n, we generate an ensemble of artificial syntactic dependency structures by exhaustive sampling for $n \leq n^* = 10$ and random sampling for $n > n^*$. These artificial syntactic dependency trees are only constrained by the definition of the different classes. They are thus free from any memory constraint other than the ones the different classes of grammars may, perhaps, impose indirectly. Strikingly, while previous work on natural languages has shown that dependency lengths are considerably below what would be expected by a random baseline without memory constraints (Ferrer-i-Cancho, 2004; Ferrer-i-Cancho & Liu, 2014), we still observe a drop in dependency lengths for randomly generated, mildly non-projective structures that supposedly abstract away from cognitive limitations. Our current findings show that memory limitations have permeated current linguistic conceptions of grammar, suggesting that it may not be possible to adequately capture our unbounded capacity for language without incorporating cognitive constraints into the grammar formalism.

Beyond upending longheld assumptions about the nature of human linguistic productivity, our findings also have key implications for debates on how children learn language and how language evolved. Whereas a common assumption in the acquisition literature is that children come to the task of language learning with built-in linguistic constraints on what they learn (Gold, 1967; Pinker, 2003), our results suggest that language-specific constraints may not be needed and instead be replaced by general cognitive constraints (Tomasello, 2005). The strong effects of memory on dependence distance minimization provide further support for the notion that language evolved through processes of cultural evolution shaped by the human brain (Christiansen & Chater, 2008), rather than the biological evolution of language-specific constraints (Pinker, 2003).

Acknowledgements

CGR is suppported by ERC (EU's Horizon 2020; FASTPARSE, 714150), MINECO (TIN2017-85160-C2-1-R, ANSWER-ASAP) and Xunta de Galicia (ED431B 2017/01 and an Oportunius program grant). RFC is supported by MINECO (TIN2017-89244-R) and AGAUR (2017SGR-856, MACDA).

References

Bod, R. (2013). *A new history of the humanities: The search for principles and patterns from antiquity to the present.* Oxford, UK: Oxford University Press.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*, 489 - 558.

Ferrer-i-Cancho, R. (2004). Euclidean distance between syntactically linked words. *Physical Review E*, *70*, 056135.

Ferrer-i-Cancho, R., & Liu, H. (2014). The risks of mixing dependency lengths from sequences of different length. *Glottotheory*, *5*, 143-155.

Gold, E. M. (1967). Language identification in the limit. *Information and Control*, *10*, 447-474.

Gómez-Rodríguez, C. (2016). Restricted non-projectivity: Coverage vs. efficiency. *Computational Linguistics*, *42*(4), 809–817.

Gómez-Rodríguez, C., & Ferrer-i-Cancho, R. (2017). Scarcity of crossing dependencies: a direct outcome of a specific constraint? *Physical Review E*, *96*, 062304.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it and how did it evolve? *Science*, *298*, 1569-1579.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars.* Oxford: Oxford University Press.

Liu, H., Xu, C., & Liang, J. (2017). Dependency distance: A new perspective on syntactic patterns in natural languages. *Physics of Life Reviews*, *21*, 171-193.

Miller, P. (2000). *Strong generative capacity: The semantics of linguistic formalism.* Cambridge, UK: Cambridge University Press.

Morrill, G. (2010). *Categorial grammar: logical syntax, semantics, and processing.* Oxford: Oxford University Press.

Nivre, J. (2005). *Dependency grammar and dependency parsing* (Tech. Rep. No. MSI 05133). Växjö University, School of Mathematics and Systems Engineering.

Pinker, S. (2003). Language as an adaptation to the cognitive niche. In M. H. Christiansen & S. Kirby (Eds.), *Language evolution* (p. 16-37). New York: Oxford University Press.

Tomasello, M. (2005). *Constructing a language. A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

REVISITING THE UNIFORMITARIAN HYPOTHESIS: CAN WE DETECT RECENT CHANGES IN THE TYPOLOGICAL FREQUENCIES OF SPEECH SOUNDS?

EITAN GROSSMAN*[1], ELAD EISEN[1], DMITRY NIKOLAEV[1,2], and STEVEN MORAN*[3]

*Corresponding Authors: eitan.grossman@mail.huji.ac.il, steven.moran@uzh.ch
[1]Department of Linguistics, Hebrew University of Jerusalem (HUJI), Jerusalem, Israel
[2]Department of Linguistics, University of Stockholm, Stockholm, Sweden
[3]Department of Comparative Linguistics, University of Zurich, Zurich, Switzerland

The Uniformitarian Hypothesis, which in essence stresses the time-independent unity of human languages, has been interpreted in a variety of ways, e.g., as a constraint on language change or as a constraint on synchronic distributions (Janda & Joseph, 2003). One version of the Uniformitarian Hypothesis, formulated explicitly as '[H]uman languages have always been pretty much the same in terms of the typological distribution of the units that compose them' (Newmeyer, 2002), is often assumed by theoretical linguists in order to infer universal (and time-independent) properties of human language directly from present-day distributions. For example, linguists have proposed cognitive explanations of word order distributions that crucially rely on Greenbergian word order typologies, e.g., Hawkins (1983). However, the possibility of inferring causes directly from present-day distributions has been called into question (Dryer, 1989; Nichols, 1992; Maslova, 2000).

Here we ask to what extent are the current cross-linguistic distributions of phonological segments the result of events of recent human history? Specifically, are cross-linguistically frequent speech sounds frequent because of language contact? The answer to this question bears on language evolution because present-day distributions may be, at least in part, due to the result of relatively recent historical events, such as colonization and globalization.

Our study is based on the comparison of three large-scale cross-linguistic databases, including BDPROTO (representing past states of human languages) (Marsico, Flavier, Verkerk, & Moran, 2018; Moran, Grossman, & Verkerk, To appear), PHOIBLE (representing contemporary distributions) (Moran & McCloy, 2019), and SEGBO (Grossman, Eisen, Nikolaev, & Moran, Submitted). SEGBO is a new database of phonological segment borrowing, which documents more than 1600 borrowed contrastive segments from 500+ borrowing languages, from 100+ families and 220+ donor languages.

To obtain a genealogically-balanced sample and robust estimates, we sample one language per phylum from both datasets 10k times and compute the relative frequencies of different phonemes in the two bootstrap samples (PHOIBLE and BDPROTO). These relative frequencies are presented in Fig. 1 for a subset of segments. To test if these shifts in frequencies are correlated with segment borrowing in general, we regressed the difference between means of bootstrap distributions of consonant frequencies on their borrowing frequency. The results are statistically significant ($p = 0.017$), pointing towards a possible role that segment borrowing has played in the evolution of segment inventories.

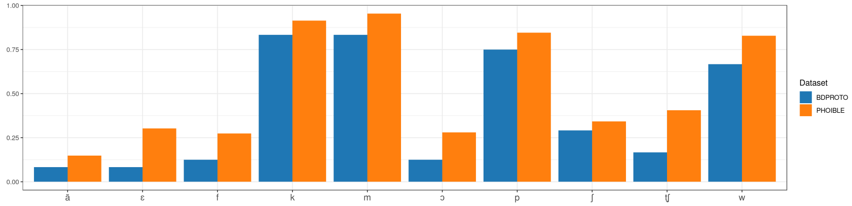


Figure 1. Mean bootstrap frequencies of a sample of segments in BDPROTO and PHOIBLE

Descriptively, we find that at a global level, some sounds are likely the result of relatively recent borrowing events, e.g., as reported by Blasi et al. (2019). In particular, the labiodental /f/ is borrowed far more frequently than any other segment (nearly 10% of the observations in the database), a fact that might point to an even later date of diffusion than previously supposed. However, different macro-areas show different patterns of borrowing. For example, languages of the Americas mostly borrowed the basic voiced stops (/b d g/), while Eurasian and Papunesian languages mostly borrowed affricates and fricatives. And African languages borrowed mostly sibilant fricatives and the voiceless bilabial stop /p/.

Each of these area-specific patterns points to a higher degree of areal specificity in pre-contact areas. Another finding is that a handful of colonial languages, which expanded in the recent past, were the major contributors of borrowed segments. Lastly, the differences in the frequencies of segments in BDPROTO and PHOIBLE correlates with the frequency of borrowing in SEGBO. Taken together, our findings point to substantive and rapid evolutionary changes in phonological segment inventories in the recent past. These findings are also relevant for linguistic theory, as they provide further evidence that present-day distributions cannot be taken as direct evidence for the naturalness of sound systems.

Acknowledgements

We thank students and colleagues at the HUJI, the University of Helsinki and the Higher School of Economics; Erich Round for a dataset on Australian languages; and Einav Levanon for his Papunesian dataset. Thanks to Swiss National Science Foundation grant/award number PCEFP1_186841 (Steven Moran, PI).

References

Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-neolithic changes in bite configuration. *Science*, *363*(6432), eaav3218.

Dryer, M. S. (1989). Large linguistic areas and language sampling. *Studies in Language*, *13*(2), 257–292.

Grossman, E., Eisen, E., Nikolaev, D., & Moran, S. (Submitted). SegBo: A database of borrowed sounds in the worlds languages.

Hawkins, J. (1983). *Word order universals.* Academic Press.

Janda, R., & Joseph, B. (2003). On language, change, and language change or, of history, linguistics, and historical linguistics. In B. Joseph & R. Janda (Eds.), *Handbook of historical linguistics* (pp. 3–180). Malden: Blackwell.

Marsico, E., Flavier, S., Verkerk, A., & Moran, S. (2018). BDPROTO: A database of phonological inventories from ancient and reconstructed languages. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018).*

Maslova, E. (2000). A dynamic approach to the verification of distributional universals. *Linguistic Typology*, *4*, 307–333.

Moran, S., Grossman, E., & Verkerk, A. (To appear). Investigating diachronic trends in phonological inventories using BDPROTO. *Language Resources & Evaluation*.

Moran, S., & McCloy, D. (Eds.). (2019). *Phoible 2.0.* Jena: Max Planck Institute for the Science of Human History.

Newmeyer, F. J. (2002). Uniformitarian assumptions and language evolution research. In A. Wray (Ed.), *The transition to language* (pp. 359–375). Oxford: Oxford University Press.

Nichols, J. (1992). *Linguistic diversity in space and time.* University of Chicago Press.

THE EMERGENCE OF COMPOSITIONAL LANGUAGES FOR NUMERIC CONCEPTS THROUGH ITERATED LEARNING IN NEURAL AGENTS

SHANGMIN GUO *[*1], YI REN[2], SERHII HAVRYLOV[2], IVAN TITOV[2], STELLA FRANK[3], and KENNY SMITH[3]

[*]Corresponding Author: sg955@cam.ac.uk
[1]Department of Computer Science and Technology, University of Cambridge, Cambridge, UK
[2]School of Informatics, University of Edinburgh, Edinburgh, UK
[3]School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK

Since first introduced by Hurford (1989), computer simulation has been an increasingly important tool in evolutionary linguistics. Recently, with the development of deep learning techniques, research in grounded language learning has also started to focus on facilitating the emergence of compositional languages without pre-defined elementary linguistic knowledge. In this work, we explore the emergence of compositional languages for numeric concepts in multi-agent communication systems. We demonstrate that compositional language for encoding numeric concepts can emerge through iterated learning in populations of deep neural network agents. However, language properties greatly depend on the input representations given to agents. We found that compositional languages only emerge if they require less iterations to be fully learnt than other non-degenerate languages for agents on a given input representation.

1. Introduction

With recent advances in deep learning (DL), it has been shown that computational agents can master a variety of complex cognitive tasks (Mnih et al., 2015; Silver et al., 2017). Recent work in grounded language learning (Hermann et al., 2017; Havrylov & Titov, 2017) applied DL techniques to enable agents to discover through learning communication protocols exhibiting language-like properties, e.g. hierarchy and compositionality. Using DL methods allow us to overcome the language pre-defining issue present in current computer simulation methods in evolutionary linguistics as in Steels (2005) and Cangelosi and Parisi (2012). The issue consists in having all basic linguistic elements (such as symbols and rules of generating phrases) to be pre-specified instead of being invented from scratch. In contrast to previous works (Mordatch & Abbeel, 2018; Cao et al., 2018) which focus on the emergence of referential signalling systems, we explore the emergent compositionality of the **non-referential** numeric concepts (which

---

[*] Work done at University of Edinburgh.

will be explained in Section 2.2) by designing a **referential** game in which agents need to transmit numerical concepts to communicate successfully.

Inspired by Kirby, Tamariz, Cornish, and Smith (2015), we model the emergence of communication protocols in dyads (i.e. the smallest possible social group of two agents) that are nodes in iterated learning chain (Kirby, 1999). We observe that iterated learning can facilitate the emergence of compositional languages for numeric concepts. However, the emergence of languages with such properties depends on the representation of numerical concepts present in the objects observed by the agents during the training. To be specific, compositional languages emerge when numeric concepts are: i) represented as a concatenation of one-hot vectors directly representing numbers; ii) implied in images of scenes featuring different number of objects. Further, we show that input representations influence the difficulty of learning a particular language by the agents, which explains the different results in case of iterated learning. For numerical concepts, we, therefore, argue that one necessary condition for the emergence of compositional languages in iterated learning is that these languages can be fully learnt [1] with less iterations for agents (especially listeners), compared with holistic languages and emergent languages from dyads.

2. Model Methods

2.1. *The Bag-Select Game*

To test whether computational agents can learn to transmit numerical concepts, we propose a referential game called as “Bag-Select” game which is briefly illustrated in Figure 2.1.

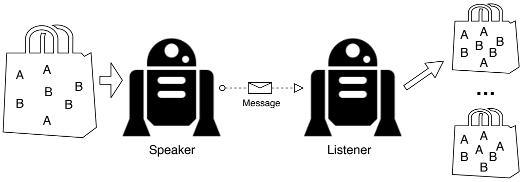


Figure 1. Sketch diagram of the Bag-Select game. The speaker observes a bag of objects of distinct types. The bag can contain a different number of objects of the specific type (here, three As and four Bs). The speaker produces a message, and the listener uses it to select the bag, that the speaker initially observed. The original bag is contained in a set among several other distinct bags, which differ only in the number of As and Bs.

In our game settings, there are two different kinds of agents: i) speaker S that observes the input b_i at the beginning of round i and then generates a message m_i; ii) listener L that receives m_i and then selects $\hat{b}_i$ among candidates c_i^k where

---

[1] A language is said to be fully learnt if: i) a speaker can always reproduce same messages as in the language given the inputs; ii) a listener could always obtain 100% accuracy given only the messages in it.

$k \in \{1, 2, \ldots, 15\}$[2] (among which one would be b_i, and the other fourteen would be uniformly sampled from the whole meaning space excluding b_i). The game only succeeds if $\hat{b}_i$ matches b_i. The speaker does not have access to the entire candidate list, only to the correct bag b_i, which implies that the number of each object type has to be encoded in the message in order to reliably succeed in the game.

2.2. *Representations of Bags*

The overall architecture of our implementation is similar to communication models proposed by Havrylov and Titov (2017). However, unlike theirs, in our game, an input b_i can be

1. **Concatenation**:a concatenation of one-hot vectors that represent numbers of each kind of objects, e.g. “2A3B” (a bag containing 2 As and 3 Bs) would be represented as $[001000; 000100]$ and “2A0B” would be represented as $[001000; 100000]$.
2. **Image**: an image containing different numbers of objects, e.g. “0A0B”, “0A2B”, “2A0B”, “2A3B”, “5A5B” would be represented as Figure 2 (a-e) respectively.
3. **Bag**: a bag of one-hot vectors that represent the quantity of different types of objects, e.g. “2A3B” and “2A0B” would be represented as $\{[01], [01], [10], [10], [10]\}$ and $\{[01], [01]\}$ respectively.

(a) “0A0B” (b) “0A2B” (c) “2A0B” (d) “2A3B” (e) “5A5B”

Figure 2. Example of an image representation of input bags that contain numerical properties. Captions under each sub-figures indicate the corresponding meaning.

As there is no specific value that can be referred to as numbers of an object in our Image and Bag representations, numeric concepts are **non-referential** in our games.

Different types of inputs require different encoders, thus we use: i) multilayer perceptron (MLP) for concatenations; ii) the convolutional neural network (CNN) which shares the same architecture of LeNet-5 proposed by LeCun, Bottou, Bengio, Haffner, et al. (1998) for images; iii) Bag-Encoder for bags.

Our bag-encoder shares almost the same architecture as the set encoder proposed by Vinyals, Bengio, and Kudlur (2015), except that we replace the softmax function in equation (5) of (Vinyals et al., 2015) with the sigmoid function. Thus, we could keep the feature representation invariant under reordering of the vectors

---

[2] In our experiments, there are always 15 candidates for listeners to choose from.

in bags, and avoid introducing normalizing bias (i.e. softmax output has to sum to one) which allows proper encoding of the numbers in the distributed representation of the bag.

To keep both meaning space and message space limited and thus analysable, there are only 2 different types of objects in our game and the maximum number of each kind of objects is 5. Therefore, the size of our Concatenation/Image dataset is 36, and the size of Bag dataset is 35 (excluding the empty bag). Messages are strings of characters of maximum length 2, where there is an available vocabulary of 10 characters.

2.3. *Iterated Learning for Neural Network Models*

We contrast two types of the population model. Following Havrylov and Titov (2017), we model dyads, pairs of agents who interact repeatedly and update their network parameters to maximise communicative success. Following Kirby et al. (2015), we contrast the communication systems that emerge in dyads with those that develop in iterated learning transmission chains. In the latter case, each generation in the chain consists of a pair of agents who are first trained on input-message pairs produced by the previous generation, then update their network parameters during communication with each other to maximise communicative success, before finally generating more data to pass to the next generation. In more detail, the model includes the following three steps:

1. **Learning phase (iterated learning only)**: During this phase, we train speaker S_t separately to reproduce same messages given the inputs, with the input-message pairs generated by S_{t-1}. For example, an input-message pair is "1A0B" → "yw", then we would train speakers to produce "yw" given the input "1A0B". To do so, we use stochastic gradient descent (SGD) (Robbins & Monro, 1951) to update parameters of S_t. Gradients are computed using the back-propagation (Rumelhart, Hinton, Williams, et al., 1988) algorithm with the cross entropy loss function between speaker's predictions and the messages generated by S_{t-1}. The number of training iterations is fixed such that predefined compositional language can be fully learnt (note that language produced by S_{t-1} is not necessarily compositional). There is no such phase in the first generation of iterated learning chain, as there are no input-message pairs for training S_1.

2. **Interaction phase**: During this phase, we train S_t and L_t agents to play the communication game using SGD. The reward is represented by the negative cross entropy between the probability distribution of the listener's prediction and the one-hot representation of the correct bag. Analogous to linguistic symbols, i.e. words, the messages transmitted between dyad should contain only discrete symbols. However, discrete messages would make learning prohibitively expensive from the computational perspective for computing the gradients would require enumeration of all possible messages. To overcome this

limitation, we use the Gumbel-softmax estimator proposed by Jang, Gu, and Poole (2016) to train our models. Besides, we set the number of iterations here to be fixed over generations, and number of iterations is obtained by pre-training a dyad to promise that it is long enough for a dyad to obtain 100% communication success rate.

3. **Transmission phase (iterated learning only)**: During this phase, we feed all b_i in the training set into S_t and sample messages m_i based on the generated probability distribution over vocabulary. This builds a dataset of input-message pairs for S_{t+1} to learn from. In addition, the number of sampled input-message pairs is $2,000$ so that they effectively reflect the distribution of all possible languages - note that since there are only 35-36 distinct input meanings to be communicated, there is no data bottleneck here, and learners will see signals for the entire space of possible meanings.

2.4. *Metrics and Evaluations*

Following Brighton and Kirby (2006), we take the topological similarity between meaning space and message space as the metric for measuring compositionality of languages, and we use Hamming distance and edit distance with respect to meaning space and message space. Equivalently, the topological similarity becomes the correlation coefficient between the Hamming distances between pairs of meanings and the edit distances between their corresponding messages. This measure captures the intuition that, in a compositional language, similar meanings will be conveyed using similar signals. We denote this measure of topological similarity as ρ; holistic (non-compositional) languages will have ρ scores around 0, a perfectly compositional language will have a ρ score of close to 1.

Additionally, we also need to measure the learning performance of new learners in order to compare the learnability of different languages, which will be introduced in Section 4. To do so, we use the accuracy of reproducing messages (both sequence-level and token-level) for speakers and accuracy of choosing the correct candidate for listeners respectively.

3. Emergence of Compositional Languages

In this section, we show that compositional languages can emerge under iterated learning, but only for the Concatenation and Image representations. As training iterated learning on deep learning models is extremely time-consuming, we report results for only one run per condition. During the exploratory phases of our research, we conducted multiple runs and found that the variance of resulting patterns of emergent languages is small, which gives us confidence that these results are representative.

To verify that iterated learning could successfully amplify the probability density of languages having high compositionality, we track the change of posterior probabilities of languages over generations. The results for the Concatenation, Image and Bag input representations are shown in the middle column of Figure 3.

As can be seen from the graphs, dyads do not converge on compositional languages under any input representation. However, in iterated learning models, the probability of languages with high compositionality ($\rho > 0.6$) keeps increasing over generations and gradually dominates all other languages, for the Concatenation and Image input representations; compositional languages do not develop in the Bag input representation. The compositional structure in the languages that emerge under the Concatenation input is clear from the example language given in Figure 3 (rightmost column), as is the absence of compositionality in the example language that develops under the Bag encoding; the final emergent language on Image representation is not perfectly compositional but contains a high degree of regularity.

4. Learnability of Compositional and Emergent Languages

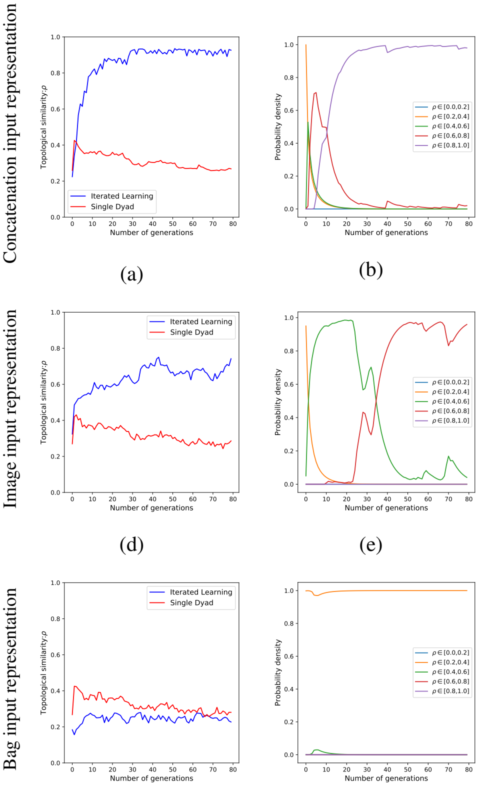
According to Kirby et al. (2015), the structure of natural languages is a trade-off between expressivity that arises during communication and compressibility that arises during learning. Meanwhile, Li and Bowling (2019) propose a hypothesis that compositional languages should be easier for listeners to learn than other less structured languages. Inspired by both of them, we hypothesise that the different effectiveness of iterated learning for different input representations observed in the above experiments is caused by different learnability of compositional languages for different input representations.

To test this hypothesis, we examine the learnability of three language types (compositional, emergent, holistic) for speakers and listeners. Our compositional test language was hand-designed and resembled the compositional languages that emerge under iterated learning in the Concatenation condition. Our holistic language was generated by randomly mapping messages from compositional languages (so that it shares same expressivity as compositional language) to inputs that constitute the whole meaning space. Our emergent test languages came from a dyad, trained to communicate as per the dyad models described above, once that dyad obtained 100% performance – as such, we would expect them to be largely holistic.

With these languages, we train speakers separately, which is illustrated in Section 2.3. At the same time, we train listeners separately to correctly complete the game with only messages in a language. For example, an input-message pair in a language is “1A0B” $\rightarrow$ “yw”, then we would train listeners to select “1A0B” among the 15 candidates after taking “yw” as input. To do so, we still take the cross entropy between the correct candidate and listener’s predicted probability distribution as the loss and apply SGD to update the parameters of listeners.

The learning curves of both listeners and speakers on different input representations are shown in Figure 4.

It is clear from Figure 4 that compositional languages require fewer training iterations than the other 2 kinds of languages in almost all the cases, with two exceptions: i) emergent languages has better learnability for listeners on the Bag



(a) (b) (d) (e) (g) (h)

	0A	1A	2A	3A	4A	5A
0B	**wq**	**yq**	**uq**	**sq**	**xq**	**zq**
1B	**wy**	**yy**	**uy**	**sy**	**xy**	**zy**
2B	**ws**	**ys**	**us**	**ss**	**xs**	**zs**
3B	**wt**	**yt**	**ut**	**st**	**xt**	**zt**
4B	**wu**	**yu**	**uu**	**su**	**xu**	**zu**
5B	**wv**	**yv**	**uv**	**sv**	**xv**	**zv**

(c)

	0A	1A	2A	3A	4A	5A
0B	**vv**	**yw**	**xu**	**xw**	**xy**	**vx**
1B	**zv**	**zz**	**xu**	**xw**	**xy**	**vx**
2B	**zv**	**zz**	**zu**	**zw**	**zy**	**zx**
3B	**uv**	**uz**	**uu**	**uw**	**uy**	**ux**
4B	**wv**	**wz**	**wu**	**ww**	**wy**	**wx**
5B	**wv**	**wz**	**wu**	**ww**	**wy**	**wx**

(f)

	0A	1A	2A	3A	4A	5A
0B		**zv**	**vz**	**vv**	**vv**	**xv**
1B	**zy**	**zu**	**zw**	**wz**	**wv**	**xw**
2B	**yz**	**uu**	**zz**	**zt**	**ww**	**wx**
3B	**yz**	**uy**	**uq**	**qz**	**tz**	**tw**
4B	**yx**	**yy**	**ur**	**qq**	**qt**	**tt**
5B	**xy**	**yr**	**ry**	**rx**	**xq**	**xt**

(i)

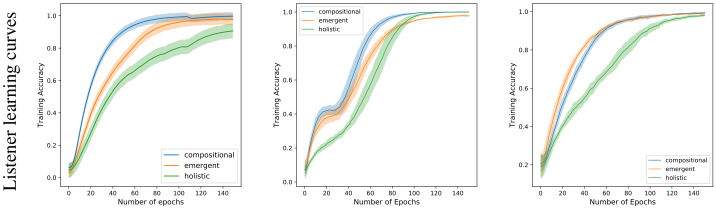
Figure 3. Experiments results on different input representations. The rows from top to bottom are results for Concatenation, Image and Bag representations respectively. The columns from left to right are: i) smoothed topological similarity (of language having greatest probability) over generations with different population models; ii) smoothed posterior probability of languages having different compositionality (ρ) over generations; iii) final emergent language facilitated by iterated learning, where the first row and first column are numbers of object "A" and "B" respectively.

representation; ii) compositional and emergent languages have almost the same learnability for speakers on the Image representation.

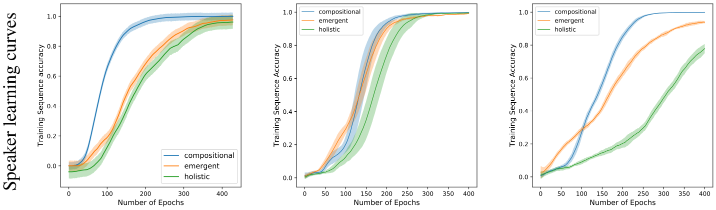
Based on the above results, considering that the topological similarity of final emergent languages given the Bag representation is much lower than Concatenation/Image representations, we argue that iterated learning will amplify the probability of compositional languages only if less training iterations are necessary for listeners to learn the compositional languages.[3] Otherwise, iterated learning

---

[3] As it is intuitive to show that compositional languages always have lower sample complexity than



(a) On Concatenation (b) On Image (c) On Bag



(d) On Concatenation (e) On Image (f) On Bag

Figure 4. Experiments results on learnability of different kinds of languages, the first row is for listeners and the second row is for speakers. input representations are given below each sub-figure. The lines are means of 10 runs with different random seeds, and the corresponding standard deviations are shown by the shadow area around the lines.

does not show lead to an increase in compositionality. Moreover, our results could also support the hypothesis that compositionality (which is an aspect of linguistic structure) emerges under the pressure of both expressivity and learnability (Smith, Tamariz, & Kirby, 2013), considering that emergent languages have better learnability on Bag representation than compositional languages; as such, those languages still represent a trade-off between learnability and expressivity, but under a slightly different learnability constraint. We are currently investigating why the Bag input encoding makes non-compositional languages more learnable.

5. Conclusion

We use the Bag-Select game to demonstrate that iterated learning leads to the emergence of compositional languages for transmitting numeric concepts. However, this result is dependent on the representations of inputs, and its effectiveness depends on that compositional languages have the optimal learnability for listeners in the communication game. While our findings confirm that structure of languages emerges under the pressure of both expressivity and learnability, at least for deep learning agents, the representation of the input representations affects on learnability and therefore on the structure of the emergent languages.

---

other non-degenerate languages and thus better learnability for speakers, we actually only need to care about learnability for listeners here, instead of both speakers and listeners as before.

References

Brighton, H., & Kirby, S. (2006). Understanding linguistic evolution by visualizing the emergence of topographic mappings. *Artificial life*, *12*(2), 229–242.

Cangelosi, A., & Parisi, D. (2012). *Simulating the evolution of language*. Springer Science & Business Media.

Cao, K., Lazaridou, A., Lanctot, M., Leibo, J. Z., Tuyls, K., & Clark, S. (2018). Emergent communication through negotiation. *arXiv preprint arXiv:1804.03980*.

Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Advances in neural information processing systems* (pp. 2149–2159). Long Beach.

Hermann, K. M., Hill, F., Green, S., Wang, F., Faulkner, R., Soyer, H., … others (2017). Grounded language learning in a simulated 3d world. *arXiv preprint arXiv:1706.06551*.

Hurford, J. R. (1989). Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, *77*(2), 187–222.

Jang, E., Gu, S., & Poole, B. (2016). Categorical reparameterization with gumbel-softmax. *arXiv preprint arXiv:1611.01144*.

Kirby, S. (1999). *Function, selection, and innateness: The emergence of language universals*. OUP Oxford.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., et al. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324.

Li, F., & Bowling, M. (2019). Ease-of-teaching and language structure from emergent communication. *arXiv preprint arXiv:1906.02403*.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., … others (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529.

Mordatch, I., & Abbeel, P. (2018). Emergence of grounded compositional language in multi-agent populations. In *Thirty-second aaai conference on artificial intelligence*.

Robbins, H., & Monro, S. (1951). A stochastic approximation method. *The annals of mathematical statistics*, 400–407.

Rumelhart, D. E., Hinton, G. E., Williams, R. J., et al. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, *5*(3), 1.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., … others (2017). Mastering the game of go without human knowledge. *Nature*, *550*(7676), 354.

Smith, K., Tamariz, M., & Kirby, S. (2013). Linguistic structure is an evolutionary trade-off between simplicity and expressivity. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 35). Berlin, Germany.

Steels, L. (2005). The emergence and evolution of linguistic structure: from lexical to grammatical communication systems. *Connection science*, *17*(3-4), 213–230.

Vinyals, O., Bengio, S., & Kudlur, M. (2015). Order matters: Sequence to sequence for sets. *arXiv preprint arXiv:1511.06391*.

A NEW HOPE (FOR COMMON GROUND)? CONVERGENCES BETWEEN BIOLINGUISTIC AND USAGE-BASED APPROACHES TO LANGUAGE AND ITS EVOLUTION

STEFAN HARTMANN[1] and MICHAEL PLEYER[*2]

[*]Corresponding Author: mpleyer@uni-koblenz.de
[1] German Department, University of Bamberg, Germany; [2] English Department,
Universität Koblenz-Landau, Campus Koblenz, Koblenz, Germany

Approaches to the study of language and its evolution are often conceptualized as belonging to one of two general theoretical frameworks: that of generative biolinguistics on the one hand (e.g. Boeckx & Grohmann 2013), and usage-based and emergentist approaches on the other (e.g. MacWhinney & O'Grady 2015). Biolinguistics traditionally stresses the language-specific genetic foundation of language and tends to adopt the theoretical commitments of generativism and the minimalist program. Usage-based and emergentist approaches, on the other hand, stress the importance of domain-general cognitive capacities, cultural factors, and interaction. For this reason they are often seen as two opposing, and seemingly irreconcilable 'camps.' In this paper, however, we argue that there have been a number of recent developments in both paradigms which suggest that biolinguistic and usage-based and emergentist approaches are actually converging on a number of key issues (also see Pleyer & Hartmann 2019). As we argue, these developments offer the potential for establishing common ground between the two approaches and offer an important step in opening up a productive dialogue on the nature of language and the factors that shape it. The fact that there are convergent trends in biolinguistics and usage-based and emergentist approaches therefore brings us closer to working towards an integrated view of language evolution, acquisition, and processing (cf. Christiansen & Chater 2016).

The convergences we observe relate to three domains specifically: a) the brain mechanisms involved in language and the degree to which the brain is specialised for processing language, b) the dynamic relationship of biology, experience and culture in the acquisition of language, and c) the interaction of cultural and biological factors in the evolution of language.

Regarding the first point, the issue of modularity and domain-specificity, many proponents of the two approaches have come to appreciate the fact that domain-general mechanisms play an important role in language. The neural systems involved in language overlap to a significant degree with other functions and tasks, and non-linguistic and linguistic activities often recruit the same neural systems. In addition, cognitive science as a whole has moved beyond simplistic conceptions of encapsulated modules, stressing instead the distributed and overlapping nature of neural and cognitive activity. As a consequence, both biolinguistics and usage-based and emergentist approaches are not asking how the modular 'language organ' (Anderson & Lightfoot 2002) came into existence but instead are taking a 'mosaic' view of the evolution of the different components that make up the language ready brain (e.g. Benítez-Burraco & Boeckx 2014).

Secondly, both approaches are beginning to integrate more sophisticated and complex views of the process of language emergence into their frameworks. Both approaches have recently been influenced by the complex adaptive systems view of language (e.g. Beckner et al. 2009, Steels 2011), evo-devo (e.g. Benítez-Burraco & Boeckx 2014; Balari & Lorenzo 2016) and niche construction theory (Laland et al., 2008; Sinha 2009), as well as developmental systems and dynamic systems theory (e.g. Overton 2015). These frameworks all stress the dynamic interplay of biology, environment, culture and interaction in ontogeny and evolution. Both biolinguistics and usage-based and emergentist approaches have therefore moved away from simplistic views of 'innateness' and towards characterizing the complex, dynamic developmental and co-evolutionary web of language evolution.

This also has direct implications for discussions of the role and relationship of biological and cultural evolution in the evolution of language. Experimental and modelling research in evolutionary linguistics has shown the importance of cultural and 'glossogenetic' factors (e.g. Steels 2011; Kirby 2017), a fact also increasingly appreciated in the biolinguistic literature (Boeckx 2017; Adger 2017). Such work can be seen as the foundation for more fruitful and focused dialogue regarding the biologically evolved aspects of language-readiness and the aspects of language shapes through processes of cultural transmission and change.

Of course, it has to be acknowledged that more traditional views of these issues still persist in certain strands of biolinguistics (e.g. Crain et al. 2016). Conversely, in usage-based and emergentist approaches, there are more radical strands much less amenable to finding common ground (e.g. Ambridge 2019). However, overall, we argue that in both paradigms, many researchers have come to embrace more complex views of these issues, enabling a more focused and productive debate over 'the evolved phenotype of language' (Balari & Lorenzo 2016) and the evolution of the language-ready brain and the interplay of cultural and biological evolution.

References

Adger, D. (2017). Restrictiveness matters. *Psychonomic Bulletin & Review, 24*(1), 138-139.

Ambridge, B. (2019). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*

Anderson, S. R., & Lightfoot, D. W. (2002). *The language organ: Linguistics as cognitive physiology*. Cambridge University Press.

Balari, S., & Lorenzo, G. (2016). Evo-devo of language and cognition. *Evolutionary Developmental Biology: A Reference Guide*, 1-14.

Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., . . . Schoenemann, T. (2009). Language is a complex adaptive system. *Language Learning, 59*(s1), 1-26.

Benítez-Burraco, A., & Boeckx, C. (2014). Universal Grammar and biological variation: an EvoDevo agenda for comparative biolinguistics. *Biological Theory, 9*(2), 122-134.

Boeckx, C. (2017), Language Evolution. In T. M. Preuss (Ed), *Human Brain Evolution* (pp. 325-339). London: Elsevier.

Boeckx, C., & Grohmann, K. K. (Eds.). (2013). *The Cambridge handbook of biolinguistics*. Cambridge University Press.

Christiansen, M. H., & Chater, N. (2016). *Creating Language: Integrating Evolution, Acquisition, and Processing*. Cambridge, Massachusetts & London, England: MIT Press.

Crain, S., Koring, L., & Thornton, R. (2017). Language acquisition from a biolinguistic perspective. *Neuroscience & Biobehavioral Reviews*, *81*, 120-149.

Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin & Review, 24*(1), 118-137.

Laland, K. N., Odling-Smee, J., & Gilbert, S. F. (2008). EvoDevo and niche construction: building bridges. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 310*(7), 549-566.

MacWhinney, B., & O'Grady, W. (2015). *The Handbook of Language Emergence*. Malden, MA: Wiley-Blackwell.

Overton, W. F. (2015). Processs, Relations, and Relational-Developmental Systems. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology developmental science. Volume 1: Theory and Method* (7th ed., pp. 9-62). Hoboken, NJ: Wiley-Blackwell.

Pleyer, M. & Hartmann, S. (2019). Constructing a Consensus on Language Evolution? Convergences and Differences Between Biolinguistic and Usage-Based Approaches. *Frontiers in Psychology 10*. doi: 10.3389/fpsyg.2019.02537.

Sinha, C. (2015). Language and other artifacts: socio-cultural dynamics of niche construction. *Frontiers in psychology*, *6*, 1601.

Steels, L. (2011). Modeling the cultural evolution of language. Physics of Life Review, 8, 339-356.

GRADUALISM IN THE EVOLUTION OF OSTENSIVE COMMUNICATION

CHRISTOPHE HEINTZ*[1] & THOM SCOTT-PHILLIPS[1,2]

[1] Department of Cognitive Science, Central European University, Budapest, Hungary
[2] Department of Anthropology, Durham University, Durham, UK
* christophe.heintz@gmail.com #istandwithceu

Most human communication is ostensive, and language use is the paradigmatic example. Here we offer a novel hypothesis about its gradual evolution in humans. We describe the graded distinctions between ostensive communication and other forms of intentional manipulation of mental states. On this basis, we show how ostensive communication could have evolved as a gradual empowerment of other social cognitive abilities to manipulate the mental states of others. We then describe the sort of social ecology in which ostensive communication is adaptive and evolutionarily stable. Specifically, we propose that cognitive processes specialised for ostensive communication will evolve only in a partner choice social ecology, where audience are able to withdraw their trust and select their informants with a high degree of possibility. We conclude with a novel suggestion about the nature of much non-human primate communication.

1. Introduction

According to many lines of argument, linguistic communication, and indeed many instances of non-linguistic communication too, such as exaggerated movements, pointing, and so on, is founded on a capacity of mind to express and recognize informative intentions, commonly called *ostensive communication* or an *interaction engine* (Sperber & Origgi, 2000; Levinson, 2006; Tomasello, 2008; Scott-Phillips, 2015). In describing this capacity, different theoretical perspectives differ in some of the detail but all agree that it is a nuanced and powerful means of social interaction, which emerges early and reliably in human development and which is grounded, one way or another, in social cognitive abilities to represent others' mental states (mindreading).

There is however a relative dearth of detail about evolutionary continuity. A common criticism is that human communicative abilities are often presented with few graded distinctions, making discussion of their gradual evolution difficult (e.g. Bar-On, 2013; Townsend et al., 2017). One approach to this problem has been to identify supposedly minimal cognitive requirements for ostensive communication (Moore, 2017a; 2017b). Others have sketched outlines of how ostensive communication could indeed evolve in a gradual manner (e.g. Sperber, 2000; Wharton, 2006; Scott-Phillips, 2015) – but these accounts lack detail, particularly about the sort of social ecology in which human communicative abilities might evolve.

Here we delineate key graded distinctions within ostensive communication, and their relevance to issues of evolutionary gradualism. We first (§2) distinguish four distinct subsets of manipulative intention (Figure 1), elaborating on each with examples. We then (§3) use this framework to describe the sort of social ecology in which the various cognitive processes involved in ostensive communication might co-evolve and be stable. We conclude (§4) that the emergence of ostensive communication in humans was driven not by the emergence of a wholly different mode of interaction, but instead by a shift in social ecology towards greater emphasis on social reputation and partner choice, which caused existing great ape social cognition to become increasingly specialised for the task of expressing and recognising informative intentions.

Note that while our analysis is focused on the production side, a complementary analysis of the comprehension side is also possible (and is not simply a mirror of the production side), but is omitted for reasons of space.

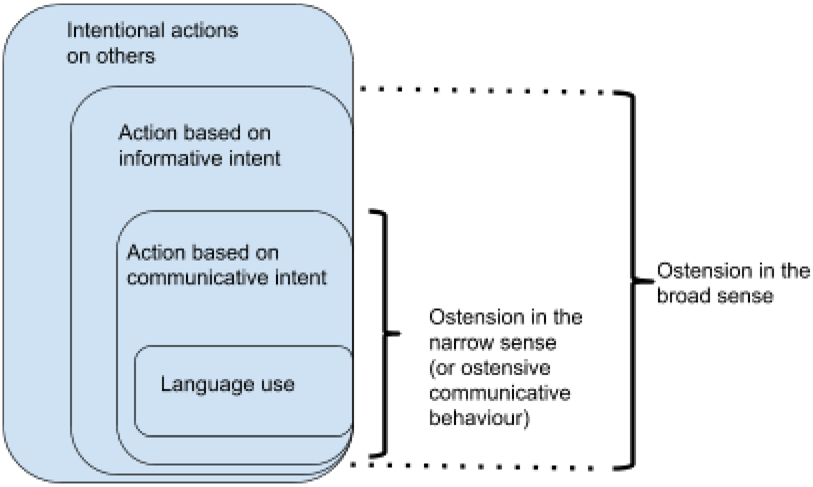


Figure 1. **Four embedded subsets of manipulative intention**. We elaborate on the contents of each in §2. See footnote 2 for the distinction between ostension in broad and narrow senses.

2. Embedded subsets of manipulative intention

Biological life is fundamentally interconnected: whatever organisms do, they cannot help but have effects on the world around them, including other organisms. In many cases effects on other organisms are merely incidental, but sometimes they are functional. Such effects are often called ‘manipulation’ (and this word is used regardless of whether the effect is beneficial to the manipulated organism; see Krebs & Dawkins, 1984; Maynard Smith & Harper, 2003). The mechanisms that produce manipulative effects are many and varied, and are often not intentional. They might instead be, for instance, only physiological, as in the case of, say, butterfly wing patterns; or chemical, as in the case of, say, quorum sensing (see Scott-Phillips et al., 2014). Here we focus only on intentional manipulation, distinguishing four embedded categories and their effects on others.

2.1. Intentional action on others

The broadest set are behaviours that are intentional and manipulative. For instance, experimental studies show how orangutan mothers will, if necessary, use their offspring as physical tools (Völter et al., 2015). Because of their small size, infants can reach food in locations that the mother cannot reach, so mothers can (and do) use them to reach the food, with the mother then consuming the food herself. Intentional manipulation can also aim at internal and mental states: some forms of startling, frightening and arousing others can all be manipulative in this way. Note, however, that they are not necessarily directly aimed at transmitting information, making others acquire beliefs, or induce any other *epistemic* change to others' mental states.

2.2. Action based on informative intent

In the second set are behaviours that are intended to inform others, and which can do so without overtly bringing attention to the informative intention itself.

An individual might dress in a smart and conservative way, as a means to inform others of her competence and professionalism yet without bringing excessive attention to oneself. Conspicuous consumption is intended to provide evidence of wealth and other markers of status. In the presence of others we might adopt a bodily posture that suggests, say, social ease and competence; and while this can be done in an overtly intentional or otherwise exaggerated way, it need not be. More generally, impression management, in which individuals present themselves in ways intended (subconsciously or otherwise) to generate and maintain a positive image in the eyes of others, but without overtly bringing attention to this informative intent, is a ubiquitous and important feature of human social life.

Such behaviour can generate a degree of shared knowledge about the actor's informative intent. However while shared knowledge is a common outcome, it is not necessarily so, and in fact in some special cases the actor might have strategic motives to actively keep her informative intent hidden or at least deniable. A criminal who plants misleading cues in a crime scene is acting on informative intent while simultaneously hiding that intent. More innocently, a dinner guest who wishes to have more wine but, recognizing it would be impolite to ask directly, might wait until her hosts' attention is elsewhere and then move her empty glass to a conspicuous location where it will, in due course, be noticed (Grosse et al., 2013). Public acts of public generosity might sometimes fall within this category also.

Looking comparatively, we take it as plausible that informative intentions exist in other primates and possibly some other species too. The key comparative questions are, in our view, whether any non-human species acts in the ways described in the next section.

2.3. Action based on communicative intent

In this third set are behaviours performed not only with an intention to inform an audience (as per §2.2) but, more than this, to make the actor's informative intent mutually known (and not just shared). Such behaviour is also known as overt intentionality or ostensive communication (but see §5).

To see the difference between this set and the one above, consider two possible ways in which Mary might satisfy her intent that Peter be informed that some berries are edible. One way is to do this is to simply eat the berries in Peter's company (without bringing any particular attention to the fact that she is doing this). In that case Mary has an informative intention which she acts on by providing evidence[1] that the berries are edible, without giving any overt evidence that she is acting on an informative intention. Instead she relies on Peter attending to her behaviour and hence drawing the inference that the berries are edible. This behaviour belongs in the second embedded subset (i.e. in §2.2). There is however an alternative. Mary might not eat the berries at all, but instead mime eating them, perhaps with exaggerated movements and while tapping her tummy. Here she has the same informative intention but provides evidence only about the intention itself, and not directly about the berries as such. Such behaviour, which makes her informative intentions mutual knowledge between Peter and Mary, belongs in the third embedded subset.

Mutual knowledge about informative intent generates meaning. This was Grice's key insight, e.g. "'*A* meant something by *x*' is (roughly) equivalent to '*A* intended the utterance of *x* to produce some effect in an audience *by means of* the recognition of this intention'" (1957, p.385, italics added). Several studies in

---

[1] Evidence independent of recognition and interpretation of Mary's informative intent.

experimental semiotics illustrate this effect in a dramatic way, by showing how behaviours produced in exaggerated ways can generate mutual knowledge about informative intent, and hence bootstrap the emergence of meaning and communication systems (e.g. Scott-Phillips et al., 2009; Newman-Norlund et al., 2009; de Ruiter et al., 2010).

However, while we have presented the distinction between this subset and the one above as categorical, it is more likely to be graded. In some cases, as with Mary's mimes or when using language, the recognition of the informative intention is essential for the observer to acquire the intended piece of information (else Peter might just think that Mary is just behaving strangely). In some other cases the recognition of the informative intention play only a small or helping role for the observer to acquire the piece of information. Suppose, for instance, that Mary eats the berries, and does so maintaining clear but not exaggerated eye contact with Peter, with some small amount of extra definition in her bodily movements. In this way she provides clear evidence that the berries are edible (as per §2.2) but also some limited evidence of her intention that Peter believe that the berries are edible (as per this section).

The possibility of graded distinctions between this set and the one above creates space for shaded differences between human ostensive communication and the social cognition of other great apes. Humans are highly competent ostensive communicators with correspondingly specialised cognitive processes (Sperber & Wilson, 2002). However we see no fundamental or inherent cognitive limitation to the expression of informative intentions in non-humans, at least in some tentative or otherwise imperfect way. In particular, informative intentions could in principle be expressed and recognised with general (not specialised) abilities of mental metarepresentation (mindreading), an ability that humans appear to share with several great ape species (Call and Tomasello, 2008; Krupenye et al., 2016). If so, then the gradual evolution of ostensive communication would involve the gradual evolution of cognitive processes dedicated to that goal.

This raises the question of why there should be any differences at all between human and non-human communication. We shall suggest that, rather than any deep cognitive limitation, social-ecological factors constrain the contexts in which action based on communicative intent is both stable and beneficial; and hence that absent such contexts, cognitive processes specialised for the expression and recognition of informative intentions are unlikely to evolve as an ordinary part of the cognitive phenotype. We develop this claim in the next section (§3), after we have described the final embedded subset, regarding language itself.

2.4. Language use

The fourth, innermost set features the same intentions as the previous set (§2.3), but is distinguished by the way in which those intentions are satisfied. Actors can make informative intentions mutually manifest (common ground) in many ways (e.g. overemphasis, mimicking, etc), including the use of culturally evolved conventions the very function of which is to help make informative intentions manifest. Pointing and nodding are examples, but the most productive means are languages: systematically structured sets of morphemes, phonemes and other constituent parts. Put simply, it is the culturally evolved function of linguistic items to help make informative intentions manifest, and hence to help make ostensive behaviour more accurate and efficacious than it otherwise would be (Origgi & Sperber, 2000; Scott-Phillips, 2017).

3. The social ecology of ostensive communication

We suggested above that the basic cognitive processes that make ostensive communication possible might be shared between great ape species, to some degree at least. At the same time, even if any non-human species does communicate in a broadly (or 'proto') ostensive way,[2] they clearly do not do so as habitually or with the same fluency as humans. Why might this be? Here we offer an answer based on differences in the social ecologies in which humans and non-human primates each live. Specifically, we suggest that the cognitive processes involved in the production of ostensive behaviours are stable only in social ecologies with high levels of partner choice i.e. those in which individuals choose between prospective partners for future social and collaborative activity.

In particular, action based on communicative intent (§2.3) will be adaptive if the audience extends, at least in a tentatively and provisional way, a presumption of cooperative intent towards the communicator. Returning to the example of Mary not eating berries but instead miming and tapping her tummy, we said that this behaviour provides evidence of her intention that Peter believe that the berries are edible – but that is true only if Peter presumes some cooperative intent on her part. If he did not then her behaviour would not motivate a search for the information that makes it worth attending the ostensive

---

[2] Within pragmatics the word ostension was first used in a precise and constrained (narrow) way, for the actions described in §2.3 and §2.4 (Sperber & Wilson, 1986/1995). Since then a sizable literature has developed, studying ostension from many perspectives including development and evolution (e.g. Gómez, 1996; Csibra, 2010; Tomasello, 2008; Moore, 2013). In the course of this progress the word has broadened in scope to sometimes include behaviours from one level further out (§2.2).

behaviour. Thus, action based on communicative intent provides benefits by increasing (massively) the communicator's capacity to manipulate mental states.

Prospective audiences can also benefit, if they are provided with relevant information – but why should that be the case? Presumptions of cooperativeness towards others are prone to exploitation and, absent other considerations, they will cause the system to collapse. And indeed such presumptions are often gamed. Lying is an attempt to exploit them, suggesting that behaviour is worthy of attention and interpretation when it is in fact not.

At the ultimate level this problem is resolved, in several species, by reputational effects (Lachmann et al., 2001; Scott-Phillips, 2008). If audiences can choose communication partners based on past reputation, communicators will tend to be honest, maintaining the stability of the system. At the proximate level the problem is resolved (in humans) by a suite of cognitive processes that filter ostensively communicated information, defending against the risk of misinformation (Sperber et al., 2010; Mercier, 2020).

Collectively known as epistemic vigilance, this suite of cognitive processes enables the stabilisation of ostensive communication in two ways. First, they impose a check on what is understood in communication, so that it might be distrusted. In other words, they differentiate comprehension from acceptance. This reduces the (expected) cost of extending a presumption of cooperativeness towards communicators, because it prevents much potentially misleading information from actually misleading. Second, epistemic vigilance allows audiences to identify misleading communicators, and hence adjust the attention and trust they are willing to grant.

Crucially, this second effect of epistemic vigilance has substantive negative consequences for unreliable communicators, who gradually lose the possibility of manipulating others' minds by means of ostensive communication – but only in a partner choice social ecology, in which individuals can gain and lose reputations, and can choose between prospective partners for future social and collaborative activity. Absent such an ecology, the cognitive processes involved in communication are highly prone to exploitation, misinformation and instability; but within it, it becomes the communicators' own interest to make it worthwhile for prospective audiences to attend to action that is based on communicative intent (Scott-Phillips, 2010; Sperber, 2013). Furthermore, the informative intent is made *mutually* manifest. This makes the communicator accountable for the relevance of what she communicates, further reinforcing the effects of partner choice (Bonalumi et al., accepted).

6. Conclusion: Some light on evolutionary gradualism

Humans have cognitive processes specialised for the core aspects of ostensive communication (e.g. Origgi & Sperber, 2000; Sperber & Wilson, 2002; Sperber et al., 2010; Csibra and Gergely, 2011; Sperber, 2013). But these various processes are all mutually dependent on one another and as such must have co-evolved, with audiences' cognitive traits constituting the social ecology of communicators', and vice-versa. A key question for language evolution is how and why this co-evolution might occur in a gradual manner.

We suggest that the answer to this question does not lie in identifying some deep cognitive limitation to non-human primates, nor in describing supposedly more minimal forms of ostensive communication (e.g. Bar-On, 2013; Moore, 2017b) – an approach which, in our view, misses the very thing that needs to be explained, namely the mentalizing inferences that are the foundation of ostensive communication in the first place. Instead, the cognitive specialisation we see in humans is largely absent in non-humans because the sort of social ecologies in which non-human primates live do not facilitate the stability of such processes. Human social ecologies involve a (much) higher degree of partner choice (Barclay & Willer, 2006; André & Baumard, 2011; Heintz et al., 2016). This is a necessary condition for the mutual stability of communicative intent and audience presumptions of cooperativeness, and hence for the gradual evolution of cognitive processes specialised for these tasks. Integration of these arguments with other ecological perspectives on the evolution of human social cognition (e.g. Tomasello et al., 2012; Sterelny, 2012; Whiten & Erdal, 2012; Moore, 2017b) is an important task for future research.

We finish with an analogy, elaborating on the idea that non-human primates might communicate in a broadly ostensive way, but without specialised cognitive processes. Consider humans swinging from trees. Our bodies are not especially well-suited to this task. We lack the specialised biological apparatus of other primates and we do not develop the relevant dispositions as an ordinary part of ontogeny. At the same time, there is no deep or fundamental barrier. Many humans can swing from trees in some ways and to some extent, and this basic ability can be refined and enhanced with training: in other words, in the right ecology. We tentatively suggest that ostensive communication in other primates might be similar: not impossible and not wholly absent, but unspecialised, somewhat disfluent, not a reliable feature of the ecology – and not part of the ordinarily developing cognitive phenotype.

References

André, J. B., & Baumard, N. (2011). The evolution of fairness in a biological market. *Evolution, 65*(5), 1447-1456.

Barclay, P., & Willer, R. (2006). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society of London B, 274*(1610), 749-753.

Bar-On, D. (2013). Origins of meaning: Must we 'go Gricean'? *Mind & Language, 28*(3), 342-375.

Bonalumi, F., Tacha, J., Scott-Phillips, T. C., & Heintz, C. (accepted). Commitment and communication: Are we committed to what we mean or what we say? *Language & Cognition*.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187-192.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25*(2), 141-168.

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society of London B, 366*(1567), 1149-1157.

De Ruiter, J. P., Noordzij, M. L., Newman-Norlund, S., Newman-Norlund, R., Hagoort, P., Levinson, S. C., & Toni, I. (2010). Exploring the cognitive infrastructure of communication. *Interaction Studies*, 11(1), 51-77.

Gómez, J. C. (1996). Ostensive behavior in the great apes: The role of eye contact. In A. Russon, S. Parker & K. Bard (Eds.), *Reaching Into Thought: The Minds Of The Great Apes* (pp. 131-151). Cambridge, Mass.: Cambridge University Press.

Grice, H. P. (1957). Meaning. *The Philosophical Review, 66*(3), 377-388.

Grosse, G., Scott-Phillips, T. C., & Tomasello, M. (2013). Three-year-olds hide their communicative intentions in appropriate contexts. *Developmental Psychology, 49*(11), 2095-2101.

Heintz, C., Karabegovic, M., & Molnar, A. (2016). The co-evolution of honesty and strategic vigilance. *Frontiers in Psychology, 7*, 1503.

Krebs, J. R. and Dawkins, R. (1984). Animal signals: Mind-reading and manipulation. In: J. R. Krebs and N. B. Davies (Eds.), *Behavioural Ecology: An evolutionary approach* (pp. 380-402). Oxford: Blackwell.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110-114.

Lachmann, M., Szamado, S., & Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences, 98*(23), 13189-13194.

Levinson, S. C. (2006). On the human 'interactional engine'. In N. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality* (pp. 39-69). Berg Publishers.

Mercier, H. (2020). *Not Born Yesterday*. Princeton University Press.

Moore, R. (2013). Evidence and interpretation in great ape gestural communication. *Humana-Mente, 24*(1): 27-51.

Moore, R. (2017a). Social cognition, stag hunts, and the evolution of language. *Biology & Philosophy, 32*(6), 797-818.

Moore, R. (2017b). Gricean communication and cognitive development. *The Philosophical Quarterly, 67*(267), 303-326.

Newman-Norlund, S. E., Noordzij, M. L., Newman-Norlund, R. D., Volman, I. A., De Ruiter, J. P., Hagoort, P., & Toni, I. (2009). Recipient design in tacit communication. *Cognition, 111*(1), 46-54.

Origgi, G., & Sperber, D. (2000). Evolution, communication and the proper function of language. In P. Carruthers & A. Chamberlain, (Eds.), *Evolution & The Human Mind: Language, modularity & social cognition* (pp. 140-169). Cambridge University Press.

Scott-Phillips, T. (2015). *Speaking Our Minds: Why human communication is different, and how language evolved to make it special*. Palgrave Macmillan.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition, 113*(2), 226-233.

Scott-Phillips, T. C. (2008). Defining biological communication. *Journal of Evolutionary Biology, 21*(2), 387-395.

Scott-Phillips, T. C. (2010). The evolution of relevance. *Cognitive Science*, *34*(4), 583-601.

Scott-Phillips, T. C., Diggle, S., Gurney, J., Ivens, A. & Popat, R. (2014). Combinatorial communication in bacteria: Implications for the origins of linguistic generativity. *PLoS One, 9*(4), e95929.

Scott-Phillips, T. C. (2017). Pragmatics and the aims of language evolution. *Psychonomic Bulletin & Review, 24*(1), 186-189.

Smith, J. M., & Harper, D. (2003). *Animal Signals*. Oxford University Press.

Sperber, D. (2013) Speakers are honest because hearers are vigilant: Reply to Kourken Michaelian. *Episteme 10*(1), 61-71.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication & Cognition*. Cambridge, MA: Harvard University Press.

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind & Language, 17*(1-2), 3-23.

Sperber, D., & Wilson, D. (2015). Beyond speaker's meaning. *Croatian Journal of Philosophy, 15(2)*(44), 117-149.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language, 25*(4), 359-393.

Sperber, D. (2000). Metarepresentations in an evolutionary perspective. In D. Sperber (ed.) *Metarepresentations: A Multidisciplinary Perspective* (pp. 117-137). Oxford University Press.

Sterelny, K. (2012). Language, gesture, skill: The co-evolutionary foundations of language. *Philosophical Transactions of the Royal Society of London B, 367*(1599), 2141-2151.

Tomasello, M. (2010). *Origins of Human Communication*. MIT Press.

Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation: The interdependence hypothesis. *Current Anthropology, 53*(6), 673-692.

Townsend, Simon W., et al. (24 authors) (2017). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biological Reviews, 92*(3), 1427-1433.

Völter, C. J., Rossano, F., & Call, J. (2015). From exploitation to cooperation: social tool use in orang-utan mother–offspring dyads. *Animal Behaviour*, *100*, 126-134.

Wharton, T. (2006). The evolution of pragmatics. *The Elsevier Encyclopaedia of Language & Linguistics*, 338-345.

Whiten, A., & Erdal, D. (2012). The human socio-cognitive niche and its evolutionary origins. *Philosophical Transactions of the Royal Society of London B, 367*(1599), 2119-2129.

COULD VOCAL COMMUNITIES OF TERRITORIAL SONGBIRDS CONSTITUTE A PSEUDO-SOCIAL SYSTEM? TESTING DIALECTS AS POTENTIAL COMMUNITY MARKERS IN SOCIAL VERSUS TERRITORIAL CONGENERIC SONGBIRD SPECIES

LAURENCE HENRY*[1], MAËL LEOUX[2], LÉA LANGE[3], ENYA QUÉRÉ[1], ISABELLE NAAS[1], ADRIAN J.F.K. CRAIG[4], and MARTINE HAUSBERGER[5]

*Corresponding Author: laurence.henry@univ-rennes1.fr
[1]Université de Rennes, Laboratoire d'Ethologie animale et humaine, UMR 6552, CNRS, Université de Caen-Normandie, France
[2]Comparative Communication and Cognition research Group, Department of Comparative Linguistics, University of Zurich
[3]CEBC – UMR 7372 – CNRS – Université de La Rochelle
[4]Department of Zoology & Entomology, Rhodes University, Grahamstown, South Africa
[5]CNRS, Laboratoire d'Ethologie animale et humaine, UMR 6552, Université de Rennes, Université de Caen-Normandie, France

Geographical variations in song features of birds are well known from ancient times and Darwin (1871) already named these variations as "provincial dialect", a parallel with human language that has been maintained and developed over the last decades (Marler & Tamura, 1962; Mundinger, 1982; Henry et al., 2015). "Dialects" or microgeographical variations where variations may occur over just a few kilometers or even a few hundred meters have been described in a variety of species (e.g. Jenkins 1978, Catchpole, 1983). These dialects result from social learning and concern both social (e.g. cacique: Feekes, 1982) and territorial (e.g. indigo bunting Paynes & Paynes 1997) species. In the latter case, birds with neighbouring territories share the same song types or local variants. Overall, vocal sharing occurs mostly between individuals living together in the same stable social group or in the same community of territorial neighbours over long periods of time. Vocal sharing/convergence is thought to convey social identity and to potentially act as a social “password” in social species (e.g. Feekes 1982, Snowdon & Hausberger 1997). Song playplack experiments show that birds tend to react more but with less aggressiveness to their own dialect. However, dialectal communities of territorial neighbours show the same features. These results have been interpreted as dear-enemies’ effects (Fisher, 1954; Briefer et al., 2008).

One other hypothesis would be that such “dialect communities” are in fact “pseudo-social” communities. Most territorial songbird species do have a social life outside the breeding season (e.g. parids: smith 1991; Ratcliffe et al., 2007) or even just outside the breeding ground (e.g. reed warbler: Catchpole; 1972; raven: Loretto et al, 2017). Moreover, they do show individual recognition of their neighbours (Molles & Vehrencamp, 2001), and do respond to their song differently according to the location of the playback (Falls & Brooks, 1975; Jaška et al. 2015). This means they have built, through repeated interactions, a representation of, and hence relationship with their closest neighbours, one of the characteristics of the social network according to Hinde (1979). In the present study, we propose that there could be a gradient between sociality and territoriality in songbird species, with dialect community of territorial neighbours being a pseudo-social system. If this is the case, we expect to find the same type of responses to the playback of own versus foreign dialects in a social species as in a close phylogenetic territorial species with dialect communities (i.e. more tolerance for own dialect).Therefore, we tested here the hypothesis that vocal dialects are “social community markers” by comparing two congeneric species of African sturnids, the territorial red-winged starling Onychognathus morio and the colonial pale-winged starling Onychognathus nabouroup. Red-winged starlings are monogamous and pairs remain together for at least three successive seasons. Breeding pairs are extremely territorial and intraspecific aggression is very common during the breeding season. However, they also spend time foraging or roosting in flocks outside the territory and during the non-breeding season. All males produce unitary whistles and birds living in adjacent territories share the same variants of the whistle types (Houdelier et al., 2012). Pale-winged starlings form also long-lived pairs but breed in colonies. Likewise they gather in flocks for foraging. Amongst the different songs of their repertoires, males produce a typical loud song type that is common to all males from the same local population but differs from one site to another throughout southern Africa. In both species, we broadcasted songs either from their own population or another distant one. The results show that the birds of both species clearly discriminated their own local variant from that of another population but also showed more attention towards their own dialect and more intolerance, like flying off, towards a foreign variant.

These results support the hypothesis that dialectal variations may constitute "social/pseudo-social" markers, reinforcing acceptance and interest between "dialect mates" and intolerance towards strangers, a feature of social communities. These findings converge with the hypothesis that dialects, both in humans and animals, may have emerged as a response to social evolution.

References

Briefer, E., Rybak, F., & Aubin. T. (2008). When to be a dear enemy: flexible acoustic relationships of neighbouring skylarks, *Alauda arvensis*. *Anim. Behav., 6,* 1319-1325.

Catchpole, C.K. (1972). A comparative study of territory in the Reed warbler (*Acrocephalus scirpaceus*) and Sedge warbler (*A. schoenobaenus*). *J. Zool., Lond., 166,* 2 1 3-23 1

Catchpole, C.K. (1983) Variation in the song of the great reed warbler Acrocephalus arundinaceus in relation to mate attraction and territorial defense. *Anim. Behav.*, 31, 1217-1225

Darwin, C. (1871). *The descent of man, and Selection in Relation to Sex.* Londres: John Murray

Falls, J. B. & Brooks, R. J. (1975). Individual recognition by song in white-throated sparrows. II. Effects of location. *Can. J. Zool.*, 53, 1412-1420.

Feekes, F. (1982). Song mimesis within colonies of *Cacicus c. cela (Leteridae, Aves).* A colonial passwords? *Z. Tierpsychol.*, 52, 119-152.

Fisher, J. 1954. Evolution and bird sociality. In J. Huxley, A. C. Hardy & E. B. Ford (Eds). *Evolution as a Process*. (pp. 71e83). London: Allen & Unwin.

Henry, L., Barbu, S., Lemasson, A., & Hausberger, M. (2015). Dialects in animals: Evidence, development and potential functions. *Anim. Behav. Cogn. 2*, 132–155. doi:10.12966/abc.05.03.2015.

Houdelier, C., Hausberger M. & Craig, A. J. F. K. (2012). Songs of two starling species: common traits versus adaptations to the social environment. *Naturwissenschaften*, 99, 1051-1062.

Hinde, R. A. (1979). *Towards understanding relationships*. London: Academic press.

Jaška, P., Pavel Linhart P. & Fuchs, R. (2015). Neighbour recognition in two sister songbird species with a simple and complex repertoire – a playback study. *J. Avian Biol.* 46, 151-158

Jenkins, P.F. (1978). Cultural transmission of song patterns and dialect development in a free-living bird population. *Anim. Behav., 26,* 50-78.

Marler, P., & Tamura, M. (1962). Song “dialects” in three populations of white-crowned sparrows. *Condor*, 64, 368-377.

Molles, L. E. & Vehrencamp, S. L. (2001). Neighbour recognition by resident males in the banded wren, Thryothorus pleurostictus, a tropical songbird with high song type sharing. *Anim. Behav.*, 61, 119-127.

Mundinger, P.C. (1982). Microgeographical and macrogeographical variation in the acquired vocalization of birds. In D.E. kroodsma & E.H. Miller (Eds.), *Acoustic communication in birds* (pp. 147-208). New York: Academic Press.

Loretto, M.C., Richard Schuster, R., Itty, C., Marchand, P., Genero, F. & Bugnyar, T. (2017). Fission-fusion dynamics over large distances in raven non-breeders. *Scientific Reports 7*: 380. DOI: 10.1038/s41598-017-00404-4

Payne, R.B., & Payne, L.L. (1997). Field observations, experimental design, and the time and place of learning bird song. In C.T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 57-84). Cambridge University Press: Cambridge.

Ratcliffe L., Mennill, D.J. & Schubert, K.A. (2007). Social dominance and fitness in black capped chickadees. In K. Otter (Ed). *Ecology and behavior of chickadees and titmice: an integrated approach*. (pp. 131–146). Oxford, UK: Oxford University Press.

Smith, S. M. (1991). *The black-capped chickadee: behavioral ecology and natural history*. Ithaca, NY: Cornell University Press

Snowdon, C.T., & Hausberger, M. (1997). *Social Influences on Vocal Development*. Cambridge: Cambridge University Press.

SHOWING AND SEEING: ENACTMENT IN BRITISH SIGN LANGUAGE CONVERSATIONS

GABRIELLE HODGE*, SANNAH GULAMANI, and KEARSY CORMIER

*Corresponding Author: g.hodge@ucl.ac.uk
Deafness Cognition & Language Research Centre, University College London, UK

It is widely accepted that investigations of enactment (non-conventional, improvised, bodily depictions of events) are integral for understanding the origins and evolution of language (see e.g. Żywiczyński, Wacewicz & Sibierska, 2018). However, there is significant disconnect in how enactment in spoken and signed languages is understood and analysed, which inhibits cross-modal comparability and investigation of the role of deaf signed languages in evolutionary theory. Here we take the position that both signers and speakers use non-conventional bodily enactment with and without more conventionalised semiotic strategies to mimetically depict the actions, utterances, thoughts and feelings of themselves, other people, animals and things (Tannen, 1989; Metzger 1995). Proficient use of enactment in deaf signed language ecologies is vital for understanding others and making oneself understood (see e.g. Cormier, Smith & Zwets, 2013; Ferrara & Johnston, 2014). Indeed, enactment is just one of several strategies for depicting in face-to-face communication, which are tightly integrated with strategies for describing and indicating (Clark, 1996; see also Ferrara & Hodge, 2018). However, unlike with spoken languages (e.g. Hakulinen & Selting, 2005), little is known about signed conversations, and the role of non-conventional semiotics during these interactions. One question is how signers use bodily enactment to visibly depict a referent while indexing other 'invisible' referents in the signing space around them. This enables signers to 'show' one referent with their body while simultaneously 'seeing' another (Winston, 1991; Engberg-Pedersen, 1993; Liddell, 2003). Here we describe how deaf signers of British Sign Language (BSL) do this during dyadic conversations, in order to highlight the coordinated complexity of depiction and indexicality within enactments occurring in everyday interactions between deaf signers of an established signed language.

Using conversations between ten deaf native and near-native signers from one geographical region (Bristol) documented in the BSL Corpus (Schembri, Fenlon, Rentelis & Cormier, 2014), we investigated: (a) who or what these signers visibly enacted with their bodies; (b) whether actions, thoughts, and/or utterances were depicted; and (c) what invisible referents were indexed via the visible enactment. For example, Figure 1 demonstrates how a deaf signer modified a manual conventionalised sign LOOK within an enactment of herself as a young child expressing a sense of surprise and wonder on seeing (for the very first time) other deaf children using signed language (represented by the white stick figures).

Manual sign: LOOK--
Visible enactment: signer looking at other children signing-------------------------------------
Invisible referent: other children signing---
Literal translation: *me looking (in wonder at the other children signing)*
Free translation: ‘I was gobsmacked looking at the other children signing.’

Figure 1. Example of visible and invisible referents (BL03F70WHC, 01:55.636-01:57.576)

We examined ~1,300 multimodal clause utterances (Enfield, 2009). Almost one-fifth of these utterances included enactment (*n*=246). Signers mainly used enactment (of varying strengths) to visibly depict their own actions, but also those of other people, and occasionally a non-human referent such as the local council. Signers also used enactment (to a much lesser extent) to depict utterances and thoughts (cf. ‘reported speech’ and ‘direct speech’). Within this set, one fifth also included invisible referents (*n*=50), mostly indexing humans other than the signer, but also organisations (e.g. the Catholic Church) and objects (e.g. a movie projector, a window). In addition to facilitating creative performance (e.g. Hodge & Ferrara, 2014) and referential cohesion (e.g. Hodge, Ferrara & Anible, 2019), these results suggest that signer’s use of enactment enables a fundamental aim of language use: to situate individual signers ‘as themselves’ within the context of the discourse, and to index the world from this perspective. We discuss these findings in light of embodied approaches to language evolution that consider multimodal strategies for indexing and depicting as a foundation of human communication (e.g. Levinson & Holler, 2014; Perlman, 2017).

References

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Cormier, K., Smith, S. & M. Zwets. (2013). Framing constructed action in British Sign Language narratives. *Journal of Pragmatics, 55*: 119-139.

Enfield, N. (2009). *The Anatomy of Meaning: Speech, Gesture, and Composite Utterances*. Cambridge: Cambridge University Press.

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: the Semantics and Morphosyntax of the Use of Space in a Visual Language*. Signum Verlag: Hamburg.

Ferrara, L. & T. Johnston. (2014). Elaborating who's what: a study of constructed action and clause structure in Auslan Australian sign language. *Australian Journal of Linguistics 34*(2): 193–215.

Ferrara, L. & G. Hodge. (2018). Language as description, indication, and depiction. *Front. Psychol. 9:716*. doi: 10.3389/fpsyg.2018.00716

Hakulinen, A., & M. Selting. (Eds.). (2005). *Syntax and Lexis in Conversation: Studies on the use of linguistic resources in talk-in-interaction*. Amsterdam: John Benjamins.

Hodge, G. & L. Ferrara. (2014). Showing the story: Enactment as performance in Auslan narratives. In L. Gawne & J. Vaughn, *Selected papers from the 44th Conference of the Australian Linguistic Society*. (pp. 372-397).

Hodge, G., Ferrara, L. & B. Anible. (2019). The semiotic diversity of doing reference in a deaf signed language. *Journal of Pragmatics, 143*: 33-53.

Levinson, S. C. & J. Holler. (2014). The origin of human multi-modal communication. *Phil. Trans. R. Soc. B 369*: 20130302. http://dx.doi.org/10.1098/rstb.2013.0302

Liddell, S. K. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. New York: Cambridge University Press.

Metzger, M. (1995). Constructed dialogue and constructed action in American Sign Language. In C. Lucas (Ed.). *Sociolinguistics in deaf communities*. Washington, DC: Gallaudet UP. (pp. 255-271).

Perlman, M. (2017). Debunking two myths against vocal origins of language: language is iconic and multimodal to the core. *Interaction Studies, 18*(3): 376-401. doi 10.1075/is.18.3.05per

Schembri, A., Fenlon, J., Rentelis, R., & Cormier, K. (2014). *British Sign Language Corpus Project: A corpus of digital video data and annotations of British Sign Language 2008-2014 (Second Edition)*. London: University College London. http://www.bslcorpusproject.org.

Tannen, D. (1989). *Talking voices*. Cambridge: Cambridge University Press.

Winston, E. (1991). Spatial referencing and cohesion in an American Sign Language text. *Sign Language Studies, 73*: 397-410.

Żywiczyński, P., Wacewicz, S. & M. Sibierska. (2018). Defining pantomime for language evolution research. *Topoi, 37*: 307-318. https://doi.org/10.1007/s11245-016-9425-9

NEUROANATOMY OF THE GREY SEAL BRAIN: BRINGING PINNIPEDS INTO THE NEUROBIOLOGICAL STUDY OF VOCAL LEARNING

NIENKE HOEKSEMA*[1,2], STELLA VILLANUEVA[3], JANINE MENGEDE[2], ANNA SALAZAR CASALS[3], ANA RUBIO GARCIA[3], BRANISLAVA CURCIC-BLAKE[4], SONJA VERNES[2,5], and ANDREA RAVIGNANI[3,6,7]

*Corresponding Author: nienke.hoeksema@mpi.nl

[1]Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[2]Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[3]Research Department, Sealcentre Pieterburen, Pieterburen, The Netherlands

[4]Department of Neuroscience, University Medical Center Groningen, Groningen, The Netherlands

[5]Donders Institute for Brain, Cognition & Behaviour, Nijmegen, The Netherlands

[6]Comparative Bioacoustics Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[7]Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium

1. Background

Vocal learning, the ability to learn novel vocalizations, is an essential part of humans' capacity for spoken language (Janik & Slater, 2000). A select group of animals – including species of bats, birds, elephants, cetaceans, and pinnipeds – also has this capability. By studying a wide range of vocal learning and non-vocal learning animals, we can increase our understanding of the neurobiological basis and evolutionary trajectory of vocal learning and human spoken language. Thus far, comparative studies have mainly centered around songbirds and non-human primates, leaving mammalian vocal learning to be understudied. Within the clade Pinnipedia, strong evidence for vocal learning has been found in both harbor seals (*Phoca vitulina*; Ralls, Fiorelli, & Gish, 1985) and grey seals (*Halichoerus grypus*; Stansbury & Janik, 2019), making them interesting targets for comparative neurobiological investigations into mammalian vocal learning. The neuroanatomy of the grey seal brain, however, has not been formally investigated.

2. Aim

The objective of the study was to conduct a first neuroanatomical investigation of the brains of grey seals to aid future comparative studies investigating the neurobiological basis of vocal learning in grey seals.

3. Methodology and Results

The brains of two female, juvenile grey seals were formalin-fixed and scanned in a 3T MRI scanner. T1- and T2-weighted image contrasts were acquired, after which the brains were dissected and photographed. A neuroanatomical atlas was created based on T2 FLAIR MR images and photographs of dissected brain slices. Moreover, a (labeled) brain template was created, as well as 3D volumetric brain models. We found that grey seal brains are larger than those of many terrestrial carnivores, and have a large cerebellum and temporal lobe, but a small olfactory area. Grey seal brains are highly convoluted, with a gyration pattern that closely matches the harbor seal brain. Building upon this work, future investigations can use diffusion tensor imaging to shed light on the neural circuits underlying vocal learning in grey seals. We are currently also exploring the expression of genes associated with vocal learning – such as FoxP2 – in collected tissue from grey seal brains via immunohistochemistry. This ongoing research will allow us to better understand the neurogenetic basis of vocal learning in grey seals.

4. Conclusion

Our study shows that the vocal learning capacity of grey seals can and should be investigated on a neurobiological level to better understand the evolution and neurobiological basis of vocal learning and human spoken language.

References

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour, 60*(1), 1–11.

Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals, Phoca vitulina. *Canadian Journal of Zoology, 63*(5), 1050–1056.

Stansbury, A. L., & Janik, V. M. (2019). Formant modification through vocal production learning in gray seals. *Current Biology, 29*(13), 1–6.

BATS AND THE COMPARATIVE NEUROBIOLOGY OF VOCAL LEARNING

NIENKE HOEKSEMA*[1,2], MAXIMILIAN WIESMANN[3,4,5], AMANDA KILIAAN[3,4,5],
PETER HAGOORT[1,6], and SONJA VERNES[2,6]

*Corresponding Author: nienke.hoeksema@mpi.nl
[1]Neurobiology of Language Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[2]Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[3]Department of Anatomy, Radboud University Medical Center, Nijmegen, The Netherlands
[4]Center for Medical Neuroscience, Donders Institute for Brain, Cognition & Behaviour, Nijmegen, The Netherlands
[5]Preclinical Imaging Center PRIME, Nijmegen, The Netherlands
[6]Donders Institute for Brain, Cognition & Behaviour, Nijmegen, The Netherlands

1. Comparative research into vocal learning and human spoken language

The capacity to communicate with one another through spoken language is a uniquely human ability. However, certain components underlying our capacity for spoken language, such as vocal production learning (from here on: vocal learning), are shared with other animals. Vocal learning is the ability to modify vocalizations based on experience (Janik & Slater, 2000). Humans use their vocal learning abilities when they are learning the phonemes and spoken words of their native language. Evidence for vocal learning in non-human animals has been found in birds, cetaceans, elephants, pinnipeds, and bats, but not in non-human primates (Fitch & Jarvis, 2013). Neurobiological research into vocal learning has mainly focused on songbirds and has taught us a great deal about the neural basis of avian vocal learning. However, as avian brains differ significantly from mammalian brains, we must also study the neurobiology of mammalian vocal learning, especially if we want to better understand the evolution and neural basis of vocal learning and human spoken language.

Bats are particularly well-suited for the study of mammalian vocal learning as bat species across the evolutionary tree show evidence of vocal learning, and numerous bat species are gregarious animals that live in social groups and employ a wide range of vocalizations (Vernes, 2017). Moreover, certain species, such as the vocal learning, pale spear-nosed bat *Phyllostomus discolor*, can be kept successfully in a laboratory setting, enabling controlled experiments (Esser, 1994; Lattenkamp, Vernes, & Wiegrebe, 2018).

2. Examining the neurobiology of bat vocal learning

Our work investigates the *P. discolor* brain using molecular and neuroimaging approaches, to facilitate testing of key hypotheses related to the neurobiology of vocal learning in a mammalian model.

In humans and songbirds, a cortical/pallial vocal motor region – the laryngeal motor cortex (LMC) in humans and the robust nucleus of the arcopallium (RA) in songbirds – is responsible for uttering voluntary vocalizations (Jürgens, 2002). It has been hypothesized that a core feature of a vocal learning brain is a direct neural connection from this area to the vocal motor neurons that control the vocal apparatus (Jürgens, 2002; Kuypers, 1958). The axon guidance genes *Robo1* and *Slit1* have been implicated in the formation of this direct connection, and in humans and songbirds, a comparatively low amount of *Slit1* expression can be found in the LMC and RA compared to the surrounding brain areas (Pfenning et al., 2014; Wang et al., 2015). In our research, we are examining the expression patterns of *Robo1* and *Slit1* across the cortex of *P. discolor* brains in order to locate a potential LMC in these vocal learning bats**.** This area can be a future target for tracing experiments to test whether a direct connection from this area to the brainstem motoneurons innervating the larynx exists in *P. discolor*.

We have also conducted magnetic resonance imaging and diffusion tensor imaging of an adult, female *P. discolor* brain. As increased connectivity between different brain regions involved in vocal communication has been found in humans compared to non-human primates, we are exploring the connectivity of the *P. discolor* brain to see if similar enhanced connectivity profiles can be observed (Kumar et al., 2016; Rilling et al. 2008).

By locating brain regions involved in vocal learning in *P. discolor* and studying the connectivity between these regions, we will increase our understanding of the different and shared neurobiological mechanisms that make animals capable of vocal learning.

References

Esser, K. H. (1994). Audio-vocal learning in a non-human mammal: The lesser spear-nosed bat Phyllostomus discolor. *NeuroReport*, *5* (14), 1718–1720.

Fitch, W. T., & Jarvis, E. D. (2013). Birdsong and Other Animal Models for Human Speech, Song, and Vocal Learning. *Language, Music, and the Brain: A Mysterious Relationship*, 499–540.

Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour, 60*(1), 1–11.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26*(2), 235–258.

Kumar, V., Croxson, P. L., & Simonyan, K. (2016). Structural Organization of the Laryngeal Motor Cortical Network and Its Implication for Evolution of Speech Production. *Journal of Neuroscience, 36*(15), 4170–4181.

Kuypers, H. G. J. (1958). Corticobulbar connexions to the pons and lower brainstem in man. *Brain, 81*, 364–388.

Lattenkamp, E. Z., Vernes, S. C., & Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *The Journal of Experimental Biology, 221*(14).

Pfenning, A. R., Hara, E., Whitney, O., Rivas, M. V, Wang, R., Roulhac, P. L., … Jarvis, E. D. (2014). Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science, 346*(6215), 1256844.

Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience, 11*(4), 426–428.

Vernes, S. C. (2017). What bats have to say about speech and language. *Psychonomic Bulletin and Review, 24*(1), 1–7.

Wang, R., Chen, C. C., Hara, E., Rivas, M. V., Roulhac, P. L., Howard, J. T., … Jarvis, E. D. (2015). Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *Journal of Comparative Neurology, 523*(6), 892–906.

EVOLUTIONARY FORCES IN THE DEVELOPMENT OF THE ENGLISH PERFECT CONSTRUCTION

MICHIO HOSAKA*[1], SHIMPEI OKUDA[2], and KAZUTOSHI SASAHARA[2, 3]

*Corresponding Author: hosaka@chs.nihon-u.ac.jp
[1]Department of English Language and Literature, Nihon University, Tokyo, Japan
[2]Graduate School of Informatics, Nagoya University, Aichi, Japan
[3]JST PRESTO, Japan

1. Introduction

Both directed selection and stochastic drift are the driving forces of biological and cultural evolution, and this is also true for language evolution. The recent argument presented by Newberry et al. (2017) is that drift cannot be rejected and stochasticity has an under-appreciated role in grammatical changes in English, such as the (ir)regularization of the past-tense verb forms.

We focus on the evolution of the English perfect construction (have/be+PP (Past Participle)) and aim to detect signatures of selection and drift working there. Although the have+PP form is used only for the perfect construction in present-day English, earlier English had a variation of have/be+PP, as below.

a. we <u>ben</u> entred into shippes bond, (Hogg & Denison, 2008)
b. For ye <u>han</u> entred into myn hous by violence (Hogg & Denison, 2008)

The decline of the BE perfect has been argued for a long time but the exact chronology and nature of the process are still in controversy.

2. Data and Methods

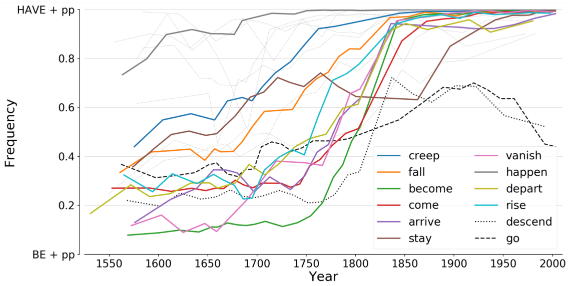
We used three English Corpora—Early English Books Online (EEBO), Corpus of Historical American English (COHA), and Google Books. Here EEBO was used to cover data from the 15th to 17th centuries (755 million words), COHA for data from the 19th to 21st centuries (406 million words), and Google Books for data of 18th century (468 billion words).

From these corpora, we computed the longitudinal frequency changes of have/be+PP forms in 19 target verbs. The target verbs were selected from the verb list of Corpus of Contemporary American English (COCA) (5764 words) if it appears in these corpora more than 200 times and the verbs' transitive usage is rare as the perfect tense and the passive voice are often indistinguishable. Moreover, we tested whether directional evolution occurred in 19 target verbs by using the Frequency Increment Test (FIT) that detects S-shaped growth. To this end, these corpora need to be comparable in size to properly set the bin size for the FIT test. Thus, we scaled the relative frequencies of the target verbs in Google Books to the frequencies of those in COHA by multiplying constant factors estimated as corresponding frequency ratios for each year as these corpora have year overlaps between 1810 and 2000. We did not scale EEBO data because there is no year overlap with other corpora and the size is approximately comparable to the size of COHA. We applied the FIT test to our corpora as in Newberry et al. (2017).

3. Results and Discussion

Fig. 1 shows that the have+PP form has become dominant over time in most verbs. For example, "become" exhibits a rapid change from 1775 to 1825. The exceptional cases of "go" and "descend" can be explained by the emergence of their adjective usages: "be gone" and "be descended." These results suggest that this auxiliary verb selection is dependent on the nature and grammatical usage of verbs. To examine whether these changes are selection or drift, we applied the FIT test as explained above. As a result, 10/19 verbs passed the FIT test ($\alpha = 0.05$), suggesting that frequency changes from be+PP to have+PP are unlikely due to random drift (i.e., directed evolution) in these verbs. Note that we cannot simply say that the verbs that did not pass the FIT test were subject to random drift; small sample size and binning procedures may affect the result (Karjus et al. 2018), thus we need further research.

We have successfully detected directional forces in the evolution of the English perfect construction. Our finding, however, needs to be further validated by making a complete list of target verbs, testing with various binning methods because several issues have been pointed out (Karjus et al. 2018), and analyzing other available corpora. If these issues are resolved, the evolutionary perspectives given here may shed a new light on grammatical changes in language evolution.



verb	FIT p-value
creep	0.001
fall	0.002
become	0.004
come	0.007
arrive	0.012
stay	0.016
vanish	0.028
happen	0.029
depart	0.045
rise	0.047
escape	0.083
ascend	0.137
prevail	0.253
arise	0.282
descend	0.437
flow	0.650
go	0.714
proceed	0.736
die	0.760

Figure 1. Relative frequencies of (have+PP) / ((have+PP) +(be+PP)) in 19 verbs (left). Result of FIT-test, in which 10 verbs that rejected the null hypothesis ($\alpha = 0.05$) are highlighted (right).

Acknowledgements

This work was supported by MEXT/JSPS KAKENHI Grant Number #4903, JP17H06379 and JP17H06383.

References

Hogg, R., & Denison, D. (2008). A history of the English language. Cambridge University Press.

Karjus, A., Blythe, R. A., Kirby, S., & Smith, K. (2018). Challenges in detecting evolutionary forces in language change using diachronic corpora. arXiv preprint arXiv:1811.01275.

Newberry, M. G., Ahern, C. A., Clark, R., & Plotkin, J. B. (2017). Detecting evolutionary forces in language change. Nature, 551(7679), 223.

THE EFFECTS OF INTERACTION ON ESTABLISHING COMMON GROUND IN COMMUNICATION

ELLIOT HUGGETT[1*], PAOLA PEÑA[2*], JUSTIN SULIK[3], and MATTHEW SPIKE[2†]

[*]Joint first author
[1]Insitute for Language, Cognition & Behaviour, CNRS, Université d'Aix-Marseille, France
[2]Centre for Language Evolution, LEL, PPLS, University of Edinburgh, UK
[3]Cognition, Values and Behaviour, Ludwig Maximilian University of Munich, Germany
[†]Corresponding Author: mspike@ed.ac.uk

Common ground and *perspective-taking* feature in most explanations of human communication. However, the relative importance of each varies across different accounts: while mentalistic approaches (e.g. Clark, 1996; Sperber & Wilson, 1986) argue that perspective taking and common ground underlie all human communication, more minimalistic accounts (e.g. Keysar et al. 2004, Pickering & Garrod, 2004) see common ground and perspective-taking as optional, controlled processes, employed only once communication has failed. In a novel signalling task Sulik & Lupyan (2018) showed that while perspective-taking alone is not sufficient for successful communication, this is partially resolved via contextual constraint. Our study extends these results to explore the circumstances under which common ground is employed in the process of communicative interaction.

Methods

Participants were recruited via Amazon Mechanical Turk in two separate experiments: 40 participants in Experiment 1, and 160 in Experiment 2.

Experiment 1: Pairs of participants took part in an interactive signalling task, in a modification of Sulik & Lupyan (2018). With participants alternating between sender and receiver roles, senders were presented with a *target* word and prompted to send any valid single English word (except the target) as a *signal* for the receiver, where the receiver then attempted to guess the original target. The sender then provided feedback to the receiver as a 1-10 rating, 10 signifying a successful guess. Each participant was allocated a subset of 6 target words randomly selected from the total set of 12, which were used across all pairs. In each round, the sender was presented with a randomly selected word from their subset; words were removed from the subset after 6 unsuccessful guesses or one successful one. The 12 English target words in the experiment were selected from published association norm corpora (Nelson et al. 2004, De Deyne, 2019) and varied in their

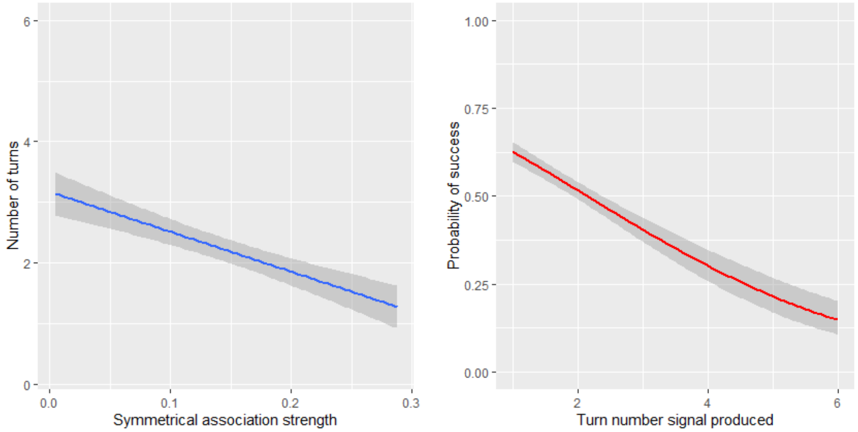


Figure 1. On the left, Exp 1: the mean number of turns to success taken for target words with different symmetrical association strengths. On the right, Exp 2: the probability of a signal which was successful in a given round in Experiment 1 being successfully mapped back to the target word.

symmetric associativity, the summed products of the targets' association strengths to other words and those words' associations strengths back to the targets.

Experiment 2: In order to test whether the participants in Experiment 1 were providing better clues over time or relying on common ground established over the interaction, naive participants were presented with the successful clues from the dyads of Experiment 1 and asked to guess the original target.

Results and discussion

Experiment 1: in a linear mixed effects regression with participant as a random effect, the number of turns to success decreased as the symmetrical association strength (SAS) increased ($\beta = 6.6, SE = 1.06, t = -6.2$, see Fig.1 left).

Experiment 2: in a linear mixed effects logistic regression with participant as a random effect, the chance of an individual successfully guessing the target increased with target difficulty ($\beta = 3.3, SE = 0.7, p \ll 0.001$) and the signal's backwards association strength ($\beta = 2.3, SE = 0.2, p \ll 0.001$), but decreased with turns to success ($\beta = -0.25, SE = 0.05, p \ll 0.001$, see Fig.1 right).

The results of experiment 1 are in line with expectations: target words which are objectively more difficult to produce signals for take more turns to be guessed. In Experiment 2, the chance of an individual successfully guessing the target is lower for signals which were successful in later rounds, even controlling for the difficulty of the target and the backwards association strength of the signal. This suggests that, rather than a gradual optimisation process where senders improve their perspective-taking and produce more informative signals, communicative success is driven by the construction of a shared context.

References

Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge university press.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The “Small World of Words” English word association norms for over 12,000 cue words. *Behavior research methods, 51*(3), 987-1006.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of personality and social psychology, 87*(3), 327.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida word association, rhyme, and word fragment norms. *Behavior Research Methods, 36*(3), 402?407.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences, 27*(2), 169-190.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Cambridge,MA: Harvard University Press.

Sulik, J., & Lupyan, G. (2018). Perspective taking in a novel signalling task: Effects of world knowledge and contextual constraint. *Journal of Experimental Psychology: General*, 147(11), 1619.

COMMUNICATIVE CONTEXT AND THE EVOLUTION OF LANGUAGE: CHANGE OVER TIME IN THE RHETORICAL STRUCTURE OF BRITISH PATENT SPECIFICATIONS, 1711 – 2011

JACK GRIEVE* and NICHOLAS GROOM

*Corresponding Author: J.Grieve@bham.ac.uk
Department of English Language and Linguistics, University of Birmingham, UK

There is a relatively long tradition of applying generalized evolutionary theoretical models to the study of language change (e.g. Croft 2000), but this research has generally investigated languages as a whole. Studies of register, genre and stylistic variation (e.g. Biber & Conrad 2019), however, have demonstrated that language varies systematically depending on the communicative contexts in which it is used. This insight is clearly relevant to evolutionary models of language change: in many ways, these communicative contexts are the cultural environments within which language evolves, much like the physical environments within which species evolve. The communicative context shapes the language used in those contexts in non-arbitrary ways, making language more suitable over time for the expression of meaning in that particular cultural domain. In this study we focus on one very specific genre – the patent specification – and offer an evolutionary account of how texts in this genre changed over time in response to cultural pressures.

Patenting is the branch of intellectual property law relating to innovations in industrial technology, and the patent specification genre lies at the heart of the entire patenting process. It is the genre in which a prospective patentee describes their invention in detail and explains why they believe it is worthy of intellectual property protection. Once it has been submitted for inspection, the specification then becomes the main focus of the patent officer's technical assessment of the inventor's claims; and if the patent application is successful, the specification finally becomes the means by which the inventor's knowledge is made available to the public, both during the period of patent protection and in perpetuity after the expiry of the patent itself.

The data for this study consist of a diachronic corpus of British patent specification texts ranging from the publication of the world's first specification in 1711 to the present day, with one patent selected at random per year. We identified systematic changes in the rhetorical structure of patent specifications over this 300-year period using the methodology of move structure analysis (Swales 1990; Biber et al 2007; Samraj 2014). In this approach, a text is seen as a sequence of 'moves', which are parts of the text that have distinct communicative functions. By conducting a move analysis for a sample of texts drawn from a given genre, it is possible to make generalisations about what types of rhetorical structures are typical of that genre, including which moves are obligatory, which are optional, and how they tend to be ordered.

We coded each of the patents in our corpus for rhetorical moves, identified through a manual analysis with inter-rater reliability testing. These moves fulfil a wide range of functions and include "filing information", "declaration of invention", and "statement of claims", for example. In this way each patent text was reduced to a sequence of moves, with each move being representing by a single orthographic character. Next, we used string edit distance techniques (Navarro 2001) to measure the dissimilarity between each temporally adjacent pair of move sequences (e.g. 1734 vs. 1735, 1735 vs. 1736). String edit distance is a relatively simple way of measuring how dissimilar two strings are from each other by counting how many changes are necessary to convert one string into another. By applying this technique to strings representing rhetorical move sequences, we were able to quantify change in the rhetorical structure in our corpus. We then plotted these string edit distances over time, thereby allowing us to visualize and identify changes in the rhetorical structure of patents (see Fig. 1). Finally, we interpreted these observed rhetorical changes from a cultural evolutionary perspective (Mesoudi 2011; Richerson & Christiansen 2013).

Our initial expectation was that the changes revealed by our data would conform either to the classic Darwinian 'phyletic gradualist' model of evolutionary change, or to the alternative 'punctuated equilibrium' model proposed by Eldredge & Gould (1972). In practice, however, our results do not fit comfortably into either of these two models, but rather combine aspects of both. Specifically, we find that the rhetorical structure of the patent specification genre is subject to constant and gradual change throughout its existence, but also that this contour of constant and gradual change is punctuated by four abrupt and dramatic shifts at key historical points in time. Accordingly, we argue that the evolution of the patent specification genre is best described, following Malmgren et al (1984), as an instance of 'punctuated gradualism'. We conclude by discussing the necessity of integrating the concept of communicative context into evolutionary theories of language change.

References

Biber, D. & Conrad, S. (2019). *Register, genre, and style*. Cambridge, UK: Cambridge University Press.

Biber, D., Connor, U., & Upton, T. A. (2007). *Discourse on the move: Using corpus analysis to describe discourse structure.* Amsterdam: John Benjamins.

Croft, W. (2000). *Explaining language change: An evolutionary approach.* Harlow, UK: Longman.

Eldredge, N., & Gould, N. E. S. J. (1972). Punctuated equilibria: an alternative to phyletic gradualism. In T.J.M. Schopf (Ed.), *Models in paleobiology* (pp. 82-115). San Francisco, CA: Freeman Cooper & Co.

Malmgren, B. A., Berggren, W. A., & Lohmann, G. P. (1984). Species formation through punctuated gradualism in planktonic foraminifera. *Science, 225(4659)*, 317-319.

Mesoudi, A. (2011). *Cultural evolution: How Darwinian theory can explain human culture and synthesize the social sciences*. Chicago, IL: University of Chicago Press.

Navarro, G. (2001). A guided tour to approximate string matching. *ACM computing surveys (CSUR), 33(1),* 31-88.

Richerson, P. J., & Christiansen, M. H. (Eds.) (2013). *Cultural evolution: Society, technology, language, and religion*. Cambridge, MA: MIT Press.

Samraj, B. (2014). Move structure. In K.P. Schneider & A. Barron (Eds.), *Pragmatics of Discourse* (pp. 385-406). Berlin: Walter de Gruyter.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings.* Cambridge, UK: Cambridge University Press.

A PHYLOGENETIC STUDY OF SINO-TIBETAN KINSHIP EVOLUTION

TING JI[1], MAY HANZHI ZHANG[2,] THOMAS E. CURRIE[3], ANDREW MEADE[4],
MARK PAGEL [4,5], and RUTH MACE[2,6*]

[*]Corresponding Author: r.mace@ucl.ac.uk
[1]Institute of Zoology, Chinese Academy of Sciences, Beijing, China
[2]Department of Anthropology, University College London, London, UK
[3]Department of Biosciences, University of Exeter, Cornwall, UK
[4]School of Biological Sciences, University of Reading, Reading, UK
[5]Santa Fe Institute, Santa Fe, New Mexico, USA
[6]School of Life Sciences, Lanzhou University, Lanzhou, China

Abstract

Descent and residence rules have long been of interest to anthropologists and biologists, as they structure populations and determine patterns of kinship, relatedness and cooperation. Despite the prevalence of patrilineal descent and patrilocal residence among extant Sino-Tibetan groups, belief in a matrilineal and matrilocal ancestry persists in China. Although some evidence on ancestral Sino-Tibetan kinship is now becoming available from both genetic and archaeological studies, the findings are contradictory[1,2]. Phylogenetic comparative methods

---

1 Key Laboratory of Animal Ecology and Conservation Biology, Centre for Computational and Evolutionary Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
2 Department of Anthropology, University College London, London WC1H 0BW, United Kingdom
3 Centre for Ecology and Conservation, Department of Bi osciences, University of Exeter, Cornwall TR10 9EZ, United Kingdom
4 School of Biological Sciences, University of Reading, Reading RG6 6UR
5 Santa Fe Institute, Santa Fe, New Mexico 87501
6 Human Evolutionary Ecology Group, School of Life Sciences, Lanzhou University, Lanzhou 730000, China
[*] e-mail: r.mace@ucl.ac.uk

(PCMs) provide an alternative to examine Sino-Tibetan kinship evolution. By mapping ethnographic data of kinship systems onto linguistic phylogenies, PCM can be used to make inferences about the pattern of cultural evolution, including ancestral states and patterns of historical change in kinship along the branches of the tree[3-5]. Here we use PCM to examine the ancestral states and trajectory of kinship evolution in Sino-Tibetan populations.

PCMs have previously been applied to questions on the evolution of kinship systems in Bantu, Indo-European, and Austronesian populations[3-5]. To this date, there is no cultural comparative study that examines Sino-Tibetan kinship evolution. Sino-Tibetan cultures are largely under-represented in all existing ethnographic databases (n=30 in Ethnographic Atlas[6,7]). We constructed a phylogenetic tree of 132 Sino-Tibetan languages using cognate data8 to represent the ancestral relationship among ethnolinguistic groups. We then coded the descent and post-marital residence norms of each group using a variety of ethnographic and historical sources[9-11]. We inferred the ancestral states of residence, descent and rates of transitions between states using Bayesian RJMCMC methods as implemented in BayesTraits[12].

Our findings show that, contrary to popular beliefs in China, matrilineal descent and matrilocal/duolocal residence are likely recent adaptations among Sino-Tibetan groups. We found strong evidence that proto-Sino-Tibetan circa. 7000 BP practised patrilineal descent and patrilocal residence. Furthermore, no unidirectional model of kinship evolution received support in our analysis, transitions from patrilineal/patrilocal to matrilineal/matrilocal systems occurred at a similar rate to transitions in the other direction. With the exception of Bodo-Koch clade, there is no statistically-significant evidence of matrilineal ancestry.
The Sino-Tibetan language family contains Sinitic languages and 200-300 Tibeto-Burman languages. It is one of the greatest language families in the world in the number of speakers. Genetic and archaeological studies have offered many alternative hypotheses regarding the relationships, geographical origin (e.g. Yellow river basin[13], the Eastern Himalayas[14], Sichuan[15]) and migration routes of the proto-Sino-Tibetans. The North-China origin hypothesis received support from two recent studies on Sino-Tibetan linguistic phylogeny16,17. However, the estimated Sino-Tibetan root age in the two studies varies (5900 years vs. 7200 years). We address the uncertainties of homeland and the exact age of Sino-Tibetan cultures by comparing the likelihoods of phylogenetic reconstructions with different candidate homeland clades fixed as the outgroup, and with different

ages of the root. Our qualitative findings are robust to uncertainties surrounding Sino-Tibetan homeland and population history.

References

1 Dong, Y. et al. Low Mitochondrial DNA Diversity in an Ancient Population from China: Insight into Social Organization at the Fujia Site. *Human biology* 87, 71-84 (2015).

2 Nelson, S. M. & Rosen-Ayalon, M. In pursuit of gender: worldwide archaeological approaches. (Rowman Altamira, 2001).

3 Opie, C., Shultz, S., Atkinson, Q. D., Currie, T. & Mace, R. Phylogenetic reconstruction of Bantu kinship challenges Main Sequence Theory of human social evolution. *Proc Natl Acad Sci* 111, 17414-17419, doi:10.1073/pnas.1415744111 (2014).

4 Fortunato, L. Reconstructing the history of marriage and residence strategies in Indo-European-speaking societies. *Hum Biol* 83, 129-135, doi:10.3378/027.083.0108 (2011).

5 Jordan, F. M., Gray, R. D., Greenhill, S. J. & Mace, R. Matrilocal residence is ancestral in Austronesian societies. *Proc Biol Sci* 276, 1957-1964, doi:10.1098/rspb.2009.0088 (2009).

6 Murdock, G. P. Ethnographic atlas. (1967).

7 Gray, J. P. A corrected ethnographic atlas. *World Cultures* 10, 24-85 (1999).

8 Peiros, I. & Starostin, S. *A comparative vocabulary of five Sino-Tibetan languages.* (The University of Melbourne, Department of Linguistics and Applied Linguistics, 1996).

9 Wang, Z. *History of nationalities in China*. China Social Science Press, Beijing (in Chinese) (1994).

10 Academia-Sinica. *中研院歷史語言研究所集刊論文類編: 民族與社會編 A Collection of Acaememia Sinica History and Language Studies Theses 1930-2000: Ethnic and Societies* Report No. 9787101062861, (中華書局 Zhonghua Publishing House, 2009).

11 Yang, S. & Ding, H. *The Study of Chinese Ethnography 中国民族志.* (Beijing: Central University for Nationalities Press, 2003).

12 Pagel, M. & Meade, A. A phylogenetic mixture model for detecting pattern-heterogeneity in gene sequence or character-state data. *Syst Biol* 53, 571-581, doi:10.1080/10635150490468675 (2004).

13 Matisoff, J. A. Sino-Tibetan linguistics: present state and future prospects. *Annual review of anthropology* 20, 469-504 (1991).

14 Blench, R. & Post, M. in *Paper from the 16th Himalayan Languages Symposium* Vol. 25 (2010).

15 van Driem, G. Language change, conjugational morphology and the Sino-Tibetan Urheimat. *Acta Linguistica Hafniensia* 26, 45-56 (1993).

16 Zhang, M., Yan, S., Pan, W. & Jin, L. Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic. *Nature* 569, 112-115, doi:10.1038/s41586-019-1153-z (2019).

17 Sagart, L. et al. Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proc Natl Acad Sci* 116, 10317-10322, doi:10.1073/pnas.1817972116 (2019).

EVOLUTIONARY RATES AND DIRECTIONS OF WORD ORDER CHANGE ACROSS MAIN AND SUBORDINATE CLAUSES IN INDO-EUROPEAN

YINGQI JING*[1,2], PAUL WIDMER[1,2], and BALTHASAR BICKEL[1,2]

*Corresponding Author: yingqi.jing@uzh.ch
[1]Department of Comparative Language Science, Center for the Interdisciplinary Study of Language Evolution, University of Zurich, Zurich, Switzerland
[2]Center for the Interdisciplinary Study of Language Evolution, University of Zurich, Zurich, Switzerland

The evolution of major word order patterns remains an open issue. One question concerns potential differences between clause types. Some researchers suggest that subordinate clauses are more conservative in word order than main clauses (Givón, 1979; Bybee, 2002), whereas others claim that subordinate clauses are more innovative (Stockwell & Minkova, 1991), or that there is no difference at all (Kroch, 1989). A second issue is the direction of word order change. Some theories posit a preference for early head recognition, hence early V positions (Hawkins, 2014). Theories of dependency length minimization tend to predict a V-medial placement (Liu, 2008; Futrell, Mahowald, & Gibson, 2015; Ferrer-i-Cancho, 2015). Predictability maximization theory supports a V-final ordering, since this configuration can maximize the predictability of the head (McDonough, Song, Hirsh-Pasek, Golinkoff, & Lannon, 2011; Ferrer-i-Cancho, 2017).

Aiming at further resolution in these debates, we use Bayesian phylogenetic inference to test hypotheses on word order evolution in the history of Indo-European. We go beyond earlier approaches and include not only basic orders (Dryer, 2013), but all observed VA/AV, VO/OV, and AO/OA orders in main and subordinate clauses in a sample of Indo-European languages, with orders ranked by rough frequency estimates as reported in descriptive grammars. We assign probabilities for each word order in specific contexts according to grammars, e.g. a 2/3 probability for VS in all-new (thetic) main clause context, a 1/2 probability in subordinate clauses etc. With these context-dependent probabilities, we can approximate the probability distribution of word orders in real utterances by re-sampling word orders from each context (500 times), so as to incorporate uncertainty into our model.

We then model the evolution of word orders as discrete states (VA/AV, VO/OV, and AO/OA) on an explicit phylogeny (Chang, Cathcart, Hall, & Garrett, 2015), assuming a Continuous-time reversible Markov Chain model of evolution and estimating transition rates with a Bayesian MCMC approach. We choose a modified implementation in `R`'s `phytools` package (Revell, 2012) which allows probability statements on states and does not enforce categorical decision. For example, the model allows a clause type in a language to be coded as having a .8 probability of OV rather than categorically as OV or VO.

Our results reveal a preferred direction of change towards AV and AO orders, but no preference in the change of the order of V and O (Figure 1). Evolutionary rates do not differ much between main and subordinate clauses, although the order of V and A tends to be slightly more stable in main clauses (mean rate of change: 0.0026) than in subordinate clauses (mean rate of change: 0.0046).

The evolutionary bias towards AV and AO is likely due to a general subject or agent first principle (Greenberg, 1963; Napoli & Sutton-Spence, 2014). The evidence for equal rates models in the placement of V and O challenges claims on the universality of trends from V-final to V-medial (Gell-Mann & Ruhlen, 2011; Maurits & Griffiths, 2014) and suggests that this placement might be less constrained than commonly assumed. Further research is needed to assess this with a higher-resolution corpus-based approach, and in other families, before we can draw firm conclusions.

Our findings also contradict the hypothesis that subordinate clauses are particularly conservative or particularly innovative. Instead, it is likely that the factors determining word order alternations might be more similar to each other in main and subordinate clauses than is commonly assumed. This also casts doubts on the notion that main clauses allow more variation than dependent clauses synchronically (Ross, 1973) and invites a less constrained view of the evolutionary dynamics of word order.

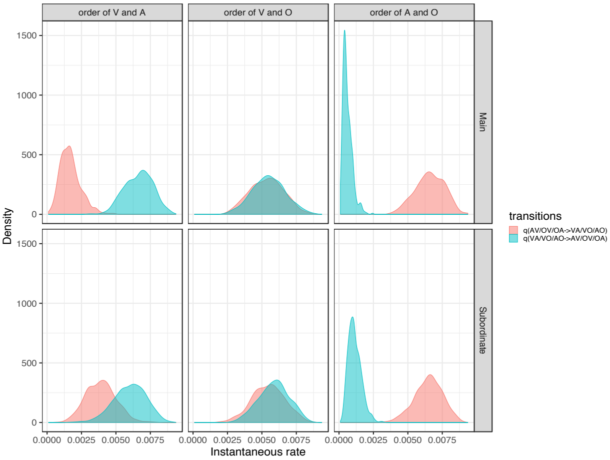


Figure 1. Posterior rate estimates

References

Bybee, J. (2002). Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. In M. N. Joan Bybee & S. Thompson (Eds.), *Complex sentences in grammar and discourse: essays in honor of Sandra A. Thompson* (pp. 1–17). Amsterdam: John Benjamins Publishing Company.

Chang, W., Cathcart, C., Hall, D., & Garrett, A. (2015). Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language*, *91*(1), 194–244.

Dryer, M. S. (2013). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology [`http://wals.info/`].

Ferrer-i-Cancho, R. (2015). The placement of the head that minimizes online memory: a complex systems approach. *Language Dynamics and Change*, *5*(1), 114–137.

Ferrer-i-Cancho, R. (2017). The placement of the head that maximizes predictability. an information theoretic approach. *Glottometrics*(39), 38–71.

Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, *112*(33), 10336–10341.

Gell-Mann, M., & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, *108*(42), 17290–17295.

Givón, T. (1979). *On understanding grammar* (Vol. 379). New York: Academic Press.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language* (p. 73-113). Cambridge, Mass.: MIT Press.

Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency.* OUP Oxford.

Kroch, A. S. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, *1*(3), 199–244.

Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, *9*(2), 159–191.

Maurits, L., & Griffiths, T. L. (2014). Tracing the roots of syntax with bayesian phylogenetics. *Proceedings of the National Academy of Sciences*, *111*(37), 13576–13581.

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (2011). An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning. *Developmental Science*, *14*(2), 181–189.

Napoli, D. J., & Sutton-Spence, R. (2014). Order of the major constituents in sign languages: implications for all language. *Frontiers in Psychology*, *5*, 376.

Revell, L. J. (2012). phytools: an r package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*(2), 217–223.

Ross, J. R. (1973). The penthouse principle and the order of constituents. In T. C. S.-S. Claudia Corum & A. Weiser (Eds.), *You take the high node and I'll take the low node* (pp. 397–422). Chicago Linguistic Society.

Stockwell, R., & Minkova, D. (1991). Subordination and word order change in the history of English. In D. Kastovsky (Ed.), *Historical English syntax* (pp. 367–408). Walter de Gruyter.

THE UNDERLYING DIMENSIONS OF SOUND-COLOR CORRESPONDENCES REVEAL TYPOLOGICALLY AND EVOLUTIONARILY GROUNDED LINGUISTIC PRIMITIVES

NIKLAS JOHANSSON*[1] and ANDREY ANIKIN[2]

*Corresponding Author: niklas.johansson@ling.lu.se
[1]Center for Languages and Literature, Lund University, Lund, Sweden
[2]Cognitive Science, Lund University, Lund, Sweden

Color plays a crucial role in language as it signals anything from danger and emotions to political affiliations, and most people have strong, sometimes synesthetic associations between colors and sounds. Sound-color mappings have therefore been extensively investigated (Spence 2011; Hamilton-Fletcher et al. 2017). However, previous studies on color symbolism and perceptual studies mostly investigated focal colors and phonemes from a few languages rather than the underlying visual and acoustic parameters. Thus, we conducted two studies, looking for the perceptual dimensions that drive sound-color correspondences and for linguistic evidence of color sound symbolism.

In study I, we tested cross-modal correspondences between each visual (luminance, hue, saturation) and each acoustic (loudness, pitch, spectral centroid, F1, F2, trill) dimensions through Implicit Associations Task experiments. Around 20 participants with varying mother tongues were recruited online for each experiment which included 16 test blocks of 16 trials each. The participants were first taught a rule associating the left arrow button with one color and sound and the right arrow button with another color and sound. They were then presented with either color or sound stimuli and were asked to press the correct arrow key as quickly as possible. In following blocks, the rule changed. Colors were sampled from the perceptually accurate *CIE-Lab* space, and the sounds were created with a formant synthesizer in order to investigate correspondences between individual visual and acoustic dimensions. Measured accuracy and reaction time were analyzed using two Bayesian mixed models: a logistic model predicting accuracy and a log-normal model predicting reaction time in correct trials. The results showed that loudness and pitch were implicitly associated with luminance and saturation. While any upward shift of spectral energy was associated with higher luminance and saturation, changing formant frequencies of synthetic vowels failed to evoke any associations, as long as the spectral centroid remained constant. We also failed to discover robust associations between the hue of isoluminant colors and any acoustic characteristic. These findings suggest that sound-color mappings on a perceptual level concern relatively simple, quantitative dimensions of colors (luminance and saturation) and sounds (auditory frequency and loudness). If the same perceptual dimensions drive sound-color symbolism in world languages, we can expect to find associations based on quantitative visual and acoustic characteristics, rather than between specific focal colors and phonemes.

In study II, we investigated eleven color terms and related concepts (red-green, yellow-blue, black-white, gray, night-day, dark-light). The color name data was gathered from 245 areally spread language families and transcribed into The International Phonetic Alphabet (IPA). Each phoneme was described acoustically using high-quality IPA recordings (Lawson et al. 2015), and average color coordinates were extracted from languages of 110 non-industrialized societies (Regier et al. 2005). The acoustic parameters (sonority, brightness, spectral centroid, F_1, F_2 and F_3 for vowels and sonority and spectral centroid for consonants) were then correlated with the color words' visual parameters (luminance and saturation). As predicted from the results of Study I, vowels with high perceived brightness, sonority and F_1 were overrepresented in names of colors with high luminance. In addition, color saturation was associated with the sonority of consonants. Thus, our findings strongly indicate that quantitative dimensions (luminance, saturation, loudness, frequency) dominate over qualitative ones (hue, vowel quality) in color sound symbolism. The results are further corroborated by reports that synesthetes and non-synesthetes (Ward et al. 2006; Moos et al. 2014), toddlers (Mondloch & Maurer 2004) and chimpanzees (Ludwig et al. 2011) prefer to map high luminance to high pitch, and that infants (Adams 1987; Skelton et al. 2017) and macaques (Xiao et al. 2011) can distinguish between high and low saturation. Furthermore, the results also aligned strikingly well with the cross-linguistic order of how color words are lexicalized (Kay & Maffi 1999). These lexicalization patterns show that the most fundamental division of the color spectrum is between light and dark colors, followed by a division between warm and cool colors, i.e. the most and least saturated colors (Witzel & Franklin 2014). Hence, there seems to be a direct link between which parameters are used for mapping sound to color iconically and which parameters influence how colors are organized in the mental lexicon.

Thus, these findings help us understand how linguistic categories evolve and develop since semantic processing seems to be affected by fundamental cross-modal associations. These sound-color associations can furthermore be linked to the increased learnability provided by iconicity, as well as evolutionary, environmental, biological and developmental constraints.

References

Adams, R. J. (1987). An evaluation of color preference in early infancy. *Infant Behavior and Development, 10*(2), 143-150.

Hamilton-Fletcher, G., Witzel, C., Reby, D., & Ward, J. (2017). Sound properties associated with equiluminant colours. *Multisensory Research*, *30*(3-5), 337-362.

Kay, P., & Maffi, L. (1999). Color appearance and the emergence and evolution of basic color lexicons. *American anthropologist*, *101*(4), 743-760.

Lawson, E., Stuart-Smith, J., Scobbie, J. M., Nakai, S., Beavan, D., Edmonds, F., Edmonds, I., Turk, A., Timmins, C., Beck, J., Esling, J., Leplatre, G., Cowen S., Barras, W. & Durham, M. (2015). Seeing Speech: an articulatory web resource for the study of Phonetics. University of Glasgow. Available at: http://www.seeingspeech.ac.uk/.

Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proceedings of the National Academy of Sciences*, *108*(51), 20661-20665.

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 133-136.

Moos, A., Smith, R., Miller, S. R., & Simmons, D. R. (2014). Cross-modal associations in synaesthesia: vowel colours in the ear of the beholder. *i-Perception*, *5*(2), 132-142.

Regier, T., Kay, P., & Cook, R. S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences*, *102*(23), 8386-8391.

Skelton, A. E., Catchpole, G., Abbott, J. T., Bosten, J. M., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sciences, 114*(21), 5545-5550.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971-995.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, *42*(2), 264-280.

Witzel, C., & Franklin, A. (2014). Do focal colors look particularly “colorful”?. *JOSA A*, *31*(4), A365-A374.

Xiao, Y., Kavanau, C., Bertin, L., & Kaplan, E. (2011). The biological basis of a universal constraint on color naming: Cone contrasts and the two-way categorization of colors. PloS one, 6(9), e24994.

POINTER EVOLUTION POINTS TO GRADUAL EVOLUTION OF HIERARCHICAL COMPLEXITY

SVERKER JOHANSSON*[1]

*Corresponding Author: sja@du.se
[1]Dalarna university, Falun, Sweden

Chomsky (e.g. 2010) and others regard unlimited Merge as the defining feature of language, that cannot evolve gradually. The neural implementation of Merge is not well understood (Rizzi 2012, Zaccarella et al 2017), but must involve something functionally equivalent to pointers in working memory. Every Merge requires two pointers, and full syntactic trees may require dozens. Other syntactic paradigms also need pointers.

Humans do hierarchies in general better than chimpanzees. Any hierarchical thinking requires nested pointers in working memory, but they are neurologically expensive and degrade with depth (Crawford et al. 2016). Humans have larger working-memory capacity than chimpanzees, which has been proposed as key to human cognitive evolution (Read 2008, Coolidge & Wynn, 2005). Gradual evolutionary growth of pointer capacity will allow gradually increasing syntactic complexity, without saltations in the underlying computational machinery. Both depth degradation and pointer capacity naturally limit Merge even in modern humans, consistent with corpus data (e.g. Karlsson 2010).

1. Can infinity evolve?

Language is commonly said to be infinite, and this is true at least in the limited sense that there are no limits to what can be said. But some popular linguistic paradigms, notably minimalism (Chomsky 1995), postulate that language is inifinite in a stronger, more literal sense, in that the language faculty can, in principle, generate an infinite number of infinitely long sentences. The generating hierarchy-building operation (Merge, in Chomsky’s case) can build trees of arbitrary depth, without limits. This kind of inifinity is sometimes invoked as an argument against the gradual evolution of the human language capacity, roughly along the following lines (e.g. Chomsky 2010, Berwick & Chomsky 2015):

1. Language is based on Merge, which is the defining feature of language.
2. Merge is unlimited.
3. Language, generated by unlimited Merge, is infinite.

4. There is no gradual path from the finite to the infinite, no possible intermediate "semi-infinite" proto-language.
    a. Having a limited Merge, and then gradually evolving an increased limit, will not get us to infinity.
5. Merge, and thus language, must have arisen in a single step. There is no such thing as "half-Merge", no intermediate stage between no Merge and full Merge (Berwick & Chomsky 2019).
6. *Conclusion*: Language did not evolve gradually.

But humans are indubitably descended from language-less ape-like ancestors. And the notion that something as complex as language could have arisen in one fell swoop, by a single super-mutation, is not tenable (e.g. Tallerman 2014, de Boer et al. 2019).

If the human language capacity did evolve in several steps, one of the points above must be incorrect. The question is which one.

2 Merge is limited

I will focus first here on assumption #2, that Merge is unlimited, and show that this is not a correct description of the actual human language capacity.

The term "Merge" *sensu stricto* is specific to Chomskyan minimalism, but regardless of linguistic paradigm there will be a need for some kind of hierarchy-handling neural machinery in the language capacity, as human languages indubitably do have some hierarchical structure. The form that the hierarchies take differ widely between paradigms, but my argument here is intended to apply over a broad range of hierarchy-handlers. I will use "Merge" here as a label, both because it is the best-known hierarchy handler, and because most proponents of the infinity argument above are Chomskyans, but I use it *sensu lato*, as a label for a generic hierarchy-handler.

From an empirical perspective, language is clearly finite. The human brain has a finite size, and the human lifespan is finite, so infinite production is obviously impossible in practice. Actual language usage shows that Merge in practice is not just limited, but limited to fairly shallow depths – in written corpus data, it is exceedingly rare to find examples of phrasal embedding more than three layers deep, and the limits appear even tighter in spoken corpora (Karlsson 2010). Similarly, the accuracy of grammaticality judgements approaches chance level as embedding depth increases beyond what is commonly used (Christiansen & MacDonald 1999).

These performance limits keep real human languages from being infinite. Languages are still very large; a few levels of embedding combined with a normal vocabulary still allows for an astronomical number of different sentences – unlimited for all practical purposes but not infinite.

However, it has been argued since Chomsky (1965) that performance is not interesting, that performance limitations just distract attention from the real underlying linguistic competence. The latter is what linguists should study, and the latter is postulated to be infinite.

But is competence, a theoretical entity that is never directly observed, actually the proper target for language *evolution* studies? Isn't it enough to account for the evolution of actual language usage?

3 How can Merge be implemented in the brain?

The neural implementation of Merge is not well understood (Rizzi 2012, Zaccarella et al. 2017). Discussions in the literature are mainly about the computational machinery, but I will focus instead on memory needs, that are non-trivial. In order to build a syntactic structure in the brain, two types of objects need to be stored in memory: (1) lexicon storage in long-term memory (LTM), and (2) syntactic nodes in working memory (WM). LTM and WM are distinct types of memory, with distinct characteristics (Norris 2017). The leaf nodes in the syntactic tree in WM must somehow refer to lexical items in LTM. Unless entire lexical items are copied into WM nodes, this must involve something functionally equivalent to *pointers* (e.g. Reilly 2003) in working memory, where a WM node contains a reference to a LTM item (Takac & Knott 2016); this would be 0-merge *sensu* Rizzi (2016).

A Merge operation will create a new object in WM, which consists of two pointers, one to each object that is merged (plus features and whatever else is stored at each node). If both pointers refer to lexical items in LTM, we have 1-merge *sensu* Rizzi (2016). The brain must be able to handle and refer to this new composite object as a single entity, a *chunk* of memory (cf. Gobet et al. 2016, Isbilen & Christiansen 2018), for purposes of further merging; chunking is a prerequisite for Merge.

For the next level of merging, it is not enough with pointers from WM to LTM; pointers from one WM location to another WM location are also needed. WM-to-WM pointers are likely neurally distinct from WM-to-LTM pointers, as the address space is different in kind. This means three different types of Merge nodes are needed in WM:

- Merging two lexical items (1-merge of Rizzi 2016).
- Merging a merged item with a lexical item (2-merge of Rizzi 2016).
- Merging two merged items (3-merge of Rizzi 2016).

All this is done in WM, and it can be noted that, while many animals do have working memory (Carruthers 2013), there is fair evidence that humans have more of it than even our closest living relatives (Read 2008, Coolidge & Wynn, 2005).

In recent works, Chomsky does note the need for working memory (or “workspace”, as he calls it), for syntactic processing, but does not discuss either implementation or limitations (e.g Chomsky et al. 2017).

4 Pointers

Pointers – having one memory location contain a reference to another memory location – are used extensively in computer programming. Biological brains are doing many operations for which a computer would use pointers. Something functionally equivalent to pointers must be neurally implemented.

4.1 Who needs pointers?

Pointers in the brain are needed as soon as the brain in one location manipulates information that is stored elsewhere in the brain. Notably, any WM computations involving LTM items will require pointers. It is rather pointless to have both types of memory, unless you have pointers as well. But both WM and LTM are widespread among non-human animals. Fish have both WM (Hughes & Blight 1999) and LTM (Lucon-Xiccato & Bisazza 2017), so presumably all vertebrates do, though most likely not identical to human memory in capacity or capabilities (Carruthers 2013). Pointers, at least WM-to-LTM, thus have an ancient origin. This means that 0-merge (Rizzi 2016) is available to all vertebrates. For 1-merge (or higher), chunking is required, for which there is evidence in e.g. rats (Fountain & Benson 2006) as well as some other mammals and birds, but negative results for fish and amphibians (Wickelgren 1979), suggesting a more limited distribution that nevertheless includes many (all?) non-primate mammals.

For 2+merge, WM-to-WM pointers are needed. Any WM operations involving hierarchical structures would be evidence of WM-to-WM pointers. There is some evidence of hierarchical cognition in non-human primates (e.g. Seyfarth et al. 2005), but it is not strongly compelling. The jury is still out on 2+merge in non-humans.

Humans, however, are hierarchical thinkers *par excellence,* to the extent that Fitch (2014) labels us “dendrophiles” for our propensity to use hierarchical thinking and impose hierarchical structure on anything and everything. Martin & Doumas (2017) propose that this general mechanism for thinking hierarchically can be repurposed for linguistic structures.

4.2 Neural implementation of pointers

Pointers are inherently difficult to handle in the brain, as memory addressing is not a matter of just storing the number of the addressed memory cell, like it is in a computer (Crawford et al. 2016). The content addressable memory in the brain requires a fundamentally different type of pointers, that are neurally quite expensive. This is particularly true when attempts are made to scale up proposed pointer models to a human-sized address space; most models do not scale well and cannot address a human-sized memory with the number of neurons available in a human brain (Blouw et al. 2016). Exact pointers are particularly vulnerable to scaling issues, whereas different types of approximate pointers fare better (Crawford et al. 2016, Legenstein et al. 2016).

The model of Crawford et al. (2016) is attractive in this context, as it has explicitly been shown in simulation to manage the full human lexicon with a reasonable number of neurons that will actually fit within the relevant brain areas. This model is based on lossy compression of information; accuracy remains adequate for single pointers, but degrades rapidly with depth when pointers are nested in recursive structures; the degradation mimics actual human performance (as opposed to theoretical competence) on multi-level embeddings.

But pointers remain expensive and consume WM fast, especially if you have hierarchical structures with multiple pointers-to-pointers.

5 The gradual evolution of limited Merge

As reviewed in the previous section, many non-human animals have LTM and WM, of limited size, as well as WM-to-LTM pointers, which implies 0-merge. Chunking, and thus 1-merge, likewise can be found in a fair range of animals (Wickelgren 1979), though the evidence for 2+merge outside the human lineage is more limited.

Contrasting this with the capacities of modern humans, we can conclude that the evolutionary changes in the human lineage, after we parted ways with the other apes, most likely include:

- Expanded WM.
- Expanded LTM (including possibly dedicated lexical storage).
- Dendrophilia.
- WM-to-WM pointers, and likely generally enhanced pointer handling.
- Node structures for 2-merge and 3-merge.
- Nested pointer handling.
- Enhanced chunking?

None of these changes need to be linguistically motivated; they are plausible components in the general human cognitive enhancement that took place concurrently (cf. Sherwood et al. 2008). There has obviously been considerable selective pressure along the human lineage for enhanced brain size and presumable enhanced cognition (Bailey & Geary 2009). This may involve expanded working memory, as there is a strong correlation between WM size and general intelligence (e.g. Colom et al. 2008).

Given the WM cost of hierarchical structures, the WM expansion will make a big difference in hierarchy handling capacities. With a small WM, you cannot build any significant tree structures even if full-blown Merge is computationally available. A larger WM invites the evolution of enhanced pointer handling, including WM-to-WM pointers and recursively nested pointers, which were pointless when WM was limited to a couple of pointers. There is also more scope for chunking in WM, when there is room for more than a few chunks. Likewise, with a larger WM, dendrophilia (Fitch 2014) starts making sense.

Speculating a bit, we get a proto-human who has a fair-sized WM, a basic set of pointer operations, including WM-to-WM pointers, and a general Merge-like operation that can do all the merge levels of Rizzi (2016). Such a proto-human can combine lexical items into two-word phrases, and can combine two phrases into a composite utterance. Dedicated syntactic machinery such as feature-checking is still missing, leading to a rather anarchic proto-language, but the basic hierarchical structure is there. Pointers remain limited in both address space and nesting depth, imposing limits on both lexicon size and tree size. But both limits can be relaxed simply by gradually adding more neurons to the pointer machinery.

If there is selective pressure towards more expressive language, with more complex syntax and enhanced narrative capacities, this can be dealt with gradually in at least three ways:

- Further WM expansion.
- Pointer expansion as above, to handle both a human-sized address space and the corpus-attested (Karlsson 2010) nesting depth.
- The general Merge operation already available may be augmented with language-specific add-ons, gradually adding all the operations that modern Merge does beyond the actual merging (feature-checking etc.), in order to make linguistic processing more precise and efficient.

The end result is modern humans, with a modern human language capacity. Merge, or whatever hierarchical operation is actually running the human language capacity, has gradually been enhanced to its modern full-fledged form, with no infinity paradoxes blocking the way.

References

Berwick, Robert C. & Chomsky, Noam. (2015) *Why only us?* Cambridge, MA: MIT Press.

Berwick, Robert C. & Chomsky, Noam. (2019) All or nothing: No half-Merge and the evolution of syntax. *PLOS Biology* 17:e3000539.

Bailey, D.H., Geary, D.C. (2009) Hominid Brain Evolution. *Hum Nat* **20,** 67–79.

Blouw et al. (2016) Concepts as semantic pointers: A framework and computational model. *Cognitive Science* 40:1128-1162

Carruthers, Peter (2013) Evolution of working memory. *Proc Nat Acad Sci* 110:10371-10378.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press

Chomsky, N. (2010). Some simple evo devo theses: how true might they be for language? In Richard K Larson, Viviane Déprez, & Hiroko Yamakido (Eds.), *The Evolution of Human Language. Biolinguistic Perspectives*. Cambridge: Cambridge University Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press

Chomsky, N, Á. J. Gallego, & D. Ott. (2017). *Generative grammar and the faculty of language: Insights, questions, and challenges*. Unpublished manuscript. https://ling.auf.net/lingbuzz/003507

Christiansen, Morten H. & MacDonald, Maryellen C. (1999) *Processing of recursive sentence structure: Testing predictions from a connectionist model.* http://www.academia.edu/download/39800748/Processing_of_Recursive_Sentence_Structu20151108-5695-17yb6hj.pdf (retrieved 2019-09-11)

Colom et al. (2008) Working memory and intelligence are highly related constructs, but why? *Intelligence*, 36:584-606.

Coolidge, Frederick L & Wynn, Thomas (2005) Working memory, its executive functions, and the emergence of modern thinking. *Cambridge Archaeological Journal* 15:5-26.

Crawford, Eric & Gingerich, Matthew & Eliasmith, Chris (2016) Biologically plausible, human-scale knowledge representation. *Cognitive Science* 40:782-821.

de Boer, Bart, Bill Thompson, Andrea Ravignani & Cedric Boeckx. (2019) *Evolutionary dynamics do not motivate a single-mutant thoery of human language.* bioRxiv preprint 517029 https://www.biorxiv.org/content/10.1101/517029v1

Fitch, W Tecumseh (2014) Toward a computational framework for cognitive biology: unifying approaches from cognitive neuroscience and comparative cognition. *Phys Life Reviews* 11:329-364

Fountain, Stephen B. & Benson, Don M. (2006) Chunking, rule learning, and multiple item memory in rat interleaved serial pattern learning. *Learning and Motivation,* 37:95-112.

Gobet, F. & Lloyd-Kelly, M. & Lane, P. C. R. (2016) What's in a name? The multiple meanings of 'chunk' and 'chunking', *Frontiers in Psychology* 7:102

Hughes, Roger N. & Blight, Christine M. (1999) Algorithmic behaviour and spatial memory are used by two intertidal fish species to solve the radial maze. *Animal Behaviour*, 58:601-613.

Isbilen, E S & Christiansen, M H (2018) Chunk-based memory constraints on the cultural evolution of language. *Topics in Cognitive Science* 2018:1-14

Karlsson, Fred (2010) Syntactic recursion and iteration. In Harry van der Hulst, ed., *Recursion and Human Language*. Berlin/New York: Mouton de Gruyter,

Legenstein et al. (2016) Assembly pointers for variable binding in networks of spiking neurons. q-bio.NC/1611.03698 (https://arxiv.org/abs/1611.03698)

Lucon-Xiccato, Tyrone & Bisazza, Angelo (2017) Complex maze learning by fish. *Animal Behaviour*, 125: 69-75.

Norris, Dennis (2017) Short-term memory and long-term memory are still different. *Psychological Bulletin* 9:992-1009

Martin, A. E. & Doumas, L. A. A. (2017) A mechanism for the cortical computation of hierarchical linguistic structure. *PLoS Biology* 15:e2000663.

Read, Dwight W (2008) Working memory: A cognitive limit to non-human primate recursive thinking prior to hominid evolution. *Evolutionary Psychology* 6:676-714.

Reilly, Edwin D. (2003). *Milestones in Computer Science and Information Technology*. Westport: Greenwood Press.

Rizzi, Luigi (2012) Core linguistic computations: How are they expressed in the mind/brain? *Journal of Neurolinguistics* 25:489-499.

Rizzi, Luigi. (2016) Monkey morpho-syntax and merge-based systems. *Theoretical Linguistics* 42(1-2): 139–145

Seyfarth, Robert M. & Cheney, Dorothy L. & Bergman, Thore J. (2005) Primate social cognition and the origins of language. *Trends in Cognitive Sciences*, 9: 264-266.

Sherwood, C.C., Subiaul, F. and Zawidzki, T.W. (2008), A natural history of the human mind: tracing evolutionary changes in brain and cognition. *Journal of Anatomy*, 212: 426-454.

Takac, M & Knott, A (2016) *Mechanisms for storing and accessing event representations in episodic memory, and their expression in language: A neural network model*. In Proceedings of the 38th Annual Conference of the Cognitive Science Society

Tallerman, Maggie. (2014) No syntax saltation in language evolution. *Language Sciences* 46: 207-219

Wickelgren, W. A. (1979). Chunking and consolidation: A theoretical synthesis of semantic networks, configuring in conditioning, S-R versus cognitive learning, normal forgetting, the amnesic syndrome, and the hippocampal arousal system. *Psychological Review, 86*(1), 44–60.

Zaccarella et al (2017) Building by syntax: the neural basis of minimal linguistic structures. *Cerebral Cortex* 27:411-421.

ASSESSING MEASURES OF MORPHOLOIGCAL COMPLEXITY AS PREDICTORS OF LEARNING BY NEURAL NETWORKS AND HUMANS

TAMAR JOHNSON*[1], KEXIN GAO[1], JENNIFER CULBERTSON[1], HUGH RABAGLIATI[2] and KENNY SMITH[1]

*Corresponding Author: tamarjohn@gmail.com
[1]Centre for Language Evolution, University of Edinburgh, UK
[2]Department of Psychology, University of Edinburgh, UK

Morphological paradigms differ widely across languages: some feature relatively few contrasts, and others, dozens. Under the assumption that languages evolve to maximise their learnability and that simpler systems are generally easier to learn (e.g. Chater & Vitanyi, 2003), this variation is surprising. Recent work on morphological complexity has resolved this paradox by arguing that certain features of even very large paradigms make them easy to learn and use. Specifically, Ackerman and Malouf (2013) propose an information-theoretic measure, i-complexity, which captures the extent to which forms in one part of a paradigm predict others. They contrast this measure with e-complexity, which is commonly used as a measure of morphological complexity in the literature (e.g., Bickel & Nichols, 2005) and captures the number of distinctions made by the language and the different ways to mark each grammatical function. They show that languages which differ widely in their e-complexity exhibit similarly low i-complexity; this suggests that having predictive relationships between inflections (i.e. low i-complexity) reduces the learnability challenge for learners even when the morphological paradigm makes many contrasts (i.e. has high e-complexity). Here, we test these measures of morphological complexity across two experiments, to evaluate whether i-complexity in fact influences the learnability of morphological paradigms as predicted by Ackerman & Malouf.

In Experiment 1 we tested, with recurrent neural networks (RNNs) and human participants trained on artificial languages, whether i-complexity predicts the learnability of inflectional paradigm. Learners were trained on an artificial language consisting of nine nouns divided into three noun classes and marked for three numbers: singular, dual and plural (Table 1). We created two languages, matched on e-complexity but differing in i-complexity. Only in the low i-complexity paradigm, the singular form of a word always predicts its form in dual. Learning in this experiment was staged - learners were first exposed to singular forms, then plural, then dual; staging allows us to test whether having learned the singular facilitates the later learning of the dual in the paradigm where the singular is predictive. Results show that for the RNNs, i-complexity affected the speed of learning, with the low i-complexity paradigm being learnt more rapidly; however, for human participants there was only weak evidence for this effect.

Table 1. Example paradigm for low i-complexity (left) and high i-complexity languages (right).

	Singular	*Dual*	*Plural*
1	-op	-um	-ib
2	-at	-oc	-el
3	-op	-um	-od

	Singular	*Dual*	*Plural*
1	-op	-um	-ib
2	-at	-um	-el
3	-op	-oc	-od

In Experiment 2 we manipulated both i-complexity and e-complexity to test whether 1) the effect of i-complexity found in Experiment 1 holds when learning is not staged (i.e., learners are exposed to wordforms from the entire paradigm randomly), and 2) whether e-complexity has an effect on learning. For RNNs there was no significant difference in learning the low vs. high *i-complexity* paradigms. However, there was a significant difference in learning the low vs. high *e-complexity* paradigms: neural networks trained on the low e-complexity paradigms achieved higher accuracy. This was mirrored in human participants: there was no effect of i-complexity but a significant effect of e-complexity. These results suggest that the benefits of lower i-complexity are dependent on learners' prior knowledge of predictive forms (not necessarily the case in natural language learning). By contrast lower values of e-complexity are advantageous for learning independent of when different parts of the paradigm are learned.

Taken together, these results cast doubt on Ackerman and Malouf's hypothesis that i-complexity rather than e-complexity drives the learnability of morphological paradigms. Since i-complexity and e-complexity are inversely correlated, the limits on i-complexity they identify in natural languages likely reflect how the two measures relate rather than how i-complexity shapes the evolution of morphological paradigms.

References

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, 429-464.

Bickel, B., & Nichols, J. (2005). Inflectional synthesis of the verb. *The world atlas of language structures*, 94-97.

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science?. *Trends in cognitive sciences*, 7(1), 19-22.

MINIMAL PREREQUISITS FOR PROCESSING LANGUAGE STRUCTURE: A MODEL BASED ON CHUNKING AND SEQUENCE MEMORY

ANNA JON-AND*[1,2] and JÉRÔME MICHAUD[3,4]

*Corresponding Author: anna.jon-and@su.se
[1]Centre for Cultural Evolution, Stockholm University, Stockholm, Sweden
[2]School of Humanities and Media Studies, Dalarna University, Falun, Sweden
[3]Department of Sociology, University of Uppsala, Uppsala, Sweden
[4]Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden

In this paper, we address the question of what minimal cognitive features are necessary for learning to process and extract grammatical structure from language. We build a minimalistic computational model containing only the two core features *chunking* and *sequence memory* and test its capacity to identify sentence borders and parse sentences in two artificial languages. The model has no prior linguistic knowledge and learns only by reinforcement of the identification of meaningful units. In simulations, the model turns out to be successful at its tasks, indicating that it is a good starting point for an extended model with ability to process and extract grammatical structure from larger corpora of natural language. We conclude that a model with the features chunking and sequence memory, that should in the future be complemented with the ability to establish hierarchical schemas, has the potential of describing the emergence of grammatical categories through language learning.

1. Introduction

In the search for the cognitive mechanisms underlying human language learning capacity, chunking has been identified as essential for overcoming memory constraints in online language processing (Christiansen & Chater 2016). The idea that language acquisition occurs through language use and through continuous updating of linguistic knowledge encoded as chunks or constructions (Bybee 1985, Tomasello 2003) has recently been successfully implemented in a chunk-based language acquisition model (McCauley & Christiansen 2019).

Sequence learning has also been pointed out as central for the language capacity (Bybee 2002, Christiansen et al. 2002, Frank et al. 2012). Faithful sequence representation has additionally been suggested to be uniquely human (Grant & Roberts 1976, MacDonald 1993, Roberts 2002, Ghirlanda et al 2017). Combined with chunking, sequence memory enables the storing of the sequential order of a chunk's components. The combination of these two features allows for the successive building up of a hierarchy of chunks that can support the identification of meaningful constructions in language processing.

The aim of this paper is to test whether the two features chunking and sequence memory are sufficient to extract simple grammatical structure from strings of artificial languages containing structures that are typical for natural languages. By implementing a minimal model architecture where no linguistic properties are predefined and evaluate its ability to extract sentences and grammatical structure in simple artificial languages, we aim at commenting the potential of the model to represent core features of the human language learning capacity. Furthermore, we compare a model with hierarchical chunking capacity to a model with a simpler incremental chunking capacity, in order to discuss whether and how chunking supports language learning. We also aim at discussing whether this model is a good starting point for an extended model able to process and extract grammatical structure from larger corpora of natural language.

2. Model

The task of the model is to segment a stream of incoming stimuli into meaningful units, conceptualized as sentences. The input consists in two small artificial languages with simple grammars. The first language contains one transitive verb and two nouns that can have the syntactic functions of subjects or objects. The word order of the language can be subject-verb-object or object-verb-subject, depending on how it is parsed. The first language thus consists of four sentences:

(1) noun_1 verb noun_1
(2) noun_1 verb noun_2
(3) noun_2 verb noun_1
(4) noun_2 verb noun_2

The second language is similar to the first one but increases complexity by introducing the possibility of adding a subordinate clause after each noun. The subordinate clause consists of a verb and a noun, making the language recursive. For both language conditions, sentences are repeated randomly in a string that constitutes the input to be processed by the model's learning mechanism. The input contains no cues that reveal the sentence borders. The processing of the input is performed by an associative learning mechanism in which $v(s \rightarrow b)$ is the stimulus-response association between stimulus *s* and behaviour *b* which estimates the value *v* of performing behaviour *b* when encountering stimulus *s*. As a consequence of experience, an agent learns about the value of responding with *b* to *s* according to

$$\Delta v(s \rightarrow b) = \alpha [u - v(s \rightarrow b)] ,$$

where *u* is the reinforcement value and *α* regulates the rate of learning. A target behaviour *b* is associated with every stimulus and the corresponding reinforcement value *u* is positive. Otherwise, the reinforcement value *u* is negative.

In our learning simulations, the learning mechanism perceives two elements and their internal temporal order before each decision. The first element can be atomic or complex, depending on previous chunkings. When perceiving a sequence, the mechanism has two basic possibilities of behaviours:

(i) ***Place border*:** A border is placed between the first and the second element in the pair. The model then suggests that a sentence ends where the border is placed and begins where the last border was placed.

(ii) ***Chunk*:** The two elements will then form a chunk that will constitute the first element in the next perceived sequence. If the first element is already a chunk, different kinds of chunkings may occur.:

   a. ***Right-chunk*:** The last element is chunked on the right-hand side to the first, without changing the internal structure of the first element.
   b. ***Sub-chunk:*** The last element is chunked with a sub-element in the first element, causing a restructuring of the first element. The number of sub-elements to which the last element can be chunked is determined by the structure of the first element. In a binary tree structure, chunking can occur with any element or node that is accessible from the right-hand side of the tree, illustrated in Fig. 1.

Flexible chunking (where right-chunking or any accessible sub-chunking can occur) generates binary hierarchical tree structures that can have any number of left- or right-branches. Fig. 1 illustrates possible chunkings with a complex element. In the example, the last element has four chunking possibilities. The upper cross indicates right-chunking and the three subsequent crosses indicate chunkings at increasingly lower levels.

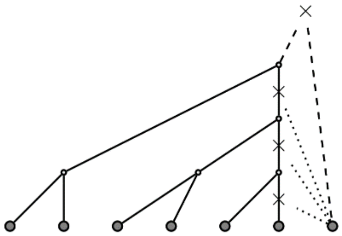


Figure 1. Illustration of possible chunkings when the first element is complex.

In order to investigate whether flexible chunking supports learning, we also test a simpler mechanism that can only right-chunk, generating purely left-branching trees. We call the two mechanisms ***flexible chunking*** and ***right-chunking***.

If a border is placed, positive reinforcement is given for correct identification of sentence boundaries and negative reinforcement for identification of incorrect boundaries. Reinforcement implies strengthening or weakening the association between the perceived sequence and the performed behaviour. This represents the concept that a language learner receives an internal or external reward for the identification of a meaningful unit i.e. a sentence. Trying to make sense of a nonsensical unit, on the other hand, can generate frustration, represented by negative reinforcement in the model. When reinforcement is given, it is also back-propagated to preceding chunkings that contributed to the successful or unsuccessful sentence identification.

Like a naïve language learner, the mechanisms in the model have no prior knowledge of grammatical structure. The mechanisms need to explore and discover on their own the chunkings that lead to correct sentence identifications and positive reinforcement.

3. Results

Results from simulations show that both the flexible chunking mechanism and the right-chunking mechanism learn to identify the sentences in the input in the two languages. As can be seen in Fig. 2, all sentences in the first, less complex language are identified and consequently pointed out after a learning process that takes the shape of an S-curve. The learning curves for the two mechanisms are very similar with the only difference that the curve of the *flexible chunking* *mechanism* has a slightly steeper S-shape. This is likely due to the larger behaviour repertoire resulting from flexible chunking, that slows down learning initially. Once succesful chunkings are identified, learning is likely faster because both productive right-chunkings and sub-chunkings are reinforced. It is not clear, however, which of the two mechanisms learns fastest.

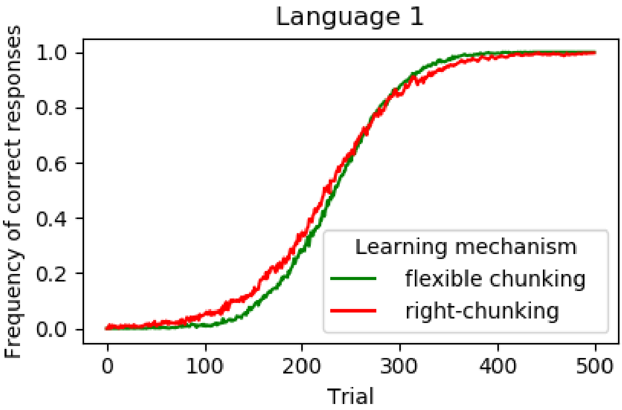


Figure 2. Simple sentences. Learning curves of the two mechanism based on sub-chunking and incremental chunking. The curves averages correct responses over 250 simulations.

After a complete learning process, both mechanisms had extracted all four grammatical sentences in the language and no agrammatical sentences. An example of grammars and parsings extracted by the mechanisms is presented in Table 1. As can be seen in Table 1, flexible chunking generates both left-branching and right-branching parsings. There seems to be no tendency for right-branching or left-branching; they are equally favoured. This is likely due to the fact that a chunking followed by a chunking or a sub-chunking are equally likely to occur and both lead to successful border placement. Once one of the two variants is tried out, it is reinforced, and the mechanism sticks to it. The right-chunking mechanism, on the other hand, generates only left-branching parsings.

Table 1. Grammars for the simple first language extracted by the two learning mechanisms

	FLEXIBLE CHUNKING	***RIGHT-CHUNKING***
Correct sentence	*Parsing*	*Parsing*
$noun_2$ verb $noun_1$	(($noun_2$ verb) $noun_1$)	(($noun_2$ verb) $noun_1$)
$noun_2$ verb $noun_2$	(($noun_2$ verb) $noun_2$)	(($noun_2$ verb) $noun_2$)
$noun_1$ verb $noun_2$	($noun_1$ (verb $noun_2$))	(($noun_1$ verb) $noun_2$)
$noun_1$ verb $noun_1$	($noun_1$ (verb $noun_1$))	(($noun_1$ verb) $noun_1$)

The second and more complex language, that involves subordinate clauses, generates a more interesting result. As seen in Fig. 3, learning to correctly identify sentences in the more complex language takes much longer for the two learning mechanisms, but it is now clear that the flexible chunking mechanism learns much faster than the right-chunking mechanism An analysis of the parsing generated by the flexible chunking mechanism offers a possible explanation to this.

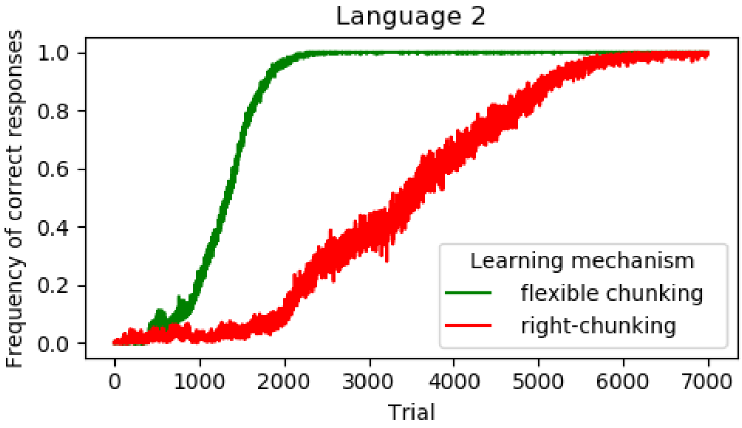


Figure 3. More complex sentences with subordinate clauses. The probability of a subordinate clause appearing after a noun is here 0.5 and the maximum number of subordinate clauses per sentence is 2. Learning curves of the two mechanisms based on flexible chunking and right-chunking. The curves average correct responses over 250 simulations.

In Table 2, only parsings generated by flexible chunking are presented. Right-chunking consistently generates left-branching parsings that do not need to be illustrated once more. As the number of possible sentences is high, some example parsings are demonstrated to illustrate the tendency that was identified.

Table 2. Examples from a grammar extracted by the flexible chunking mechanism in the more complex language with subordinate clauses.

	FLEXIBLE CHUNKING
Correct sentence	*Parsing*
noun$_2$ verb noun$_1$	((noun$_2$ verb) noun$_1$)
noun$_1$ verb noun$_2$	(noun$_1$ (verb noun$_2$))
noun$_1$ verb noun$_2$ verb noun$_2$	((noun$_1$ verb) (noun$_2$ (verb noun$_2$)))
noun$_1$ verb noun$_1$ verb noun$_1$ verb noun$_2$	(noun$_1$ (verb ((noun$_1$ verb) (noun$_1$ (verb noun$_2$)))))

As can be seen in Table 2, there is still no consequent right-branching or left-branching parsing of the recursive sentence structures with subordinate clauses. What can be observed, however, is how chunkings from shorter sentences support the building up of longer sentences.

Firstly, a chunking that is always reinforced in the identification of the shorter sentences is noun-verb. This supports the sub-chunking of a noun and a verb in subordinate clauses. In the parsing of sentences with subordinate clauses, nouns followed by a verb are always preceded by a left parenthesis, indicating that the noun and the following verb have been sub-chunked. Secondly, if a sub-chunking of the verb and a noun occurs in the parsing of a short sentence, as in the second example sentence in Table 2, where the verb is sub-chunked with noun$_2$, this sub-chunking tends to reappear in subordinate clauses. This can be seen in the last two sentences in Table 2, where this sub-chunk appears last in the parsing of both sentences.

Apart from these two tendencies, different and seemingly random parsing structures appear. It seems clear, however, that frequent chunkings in shorter sentences are reused in longer sentences. This probably explains the fast learning of the flexible chunking mechanism. The right-chunking mechanism cannot use the support from previous chunkings when learning to identify increasingly longer sentences.

4. Discussion

These first results from testing the language processing capacity of a minimal language learning model are promising for future extensions of the model. The fact that a reinforcement learning model including only the two core features *chunking* and *sequence memory* is able to learn to correctly identify sentences in small artificial languages with and without recursion is a preliminary yet powerful indication of the potential of the model. The ability of a model to extract meaningful constructions with no pre-assumptions concerning grammatical categories or rules is compatible with the idea of emergent grammatical categories in Radical Construction Grammar (Croft 2001). The comparison between the flexible chunking mechanism and the right-chunking mechanism shows that flexible chunking initially slows learning down, but if the language is complex and contains repetition of structures, learning soon becomes faster than for the right-chunking mechanism. This indicates that flexible chunking may be an important property of an incremental learning model based on chunking and sequence memory. We believe that if exposed to more complex and variable grammatical structures, such as those of natural languages, and extended with a schematizing feature, the model would most likely chose right-branching or left-branching parsing of given structures more consistently and this would likely increase the utility of flexible chunking even more.

The principle of flexible chunking is similar to that of *unsupervised data-oriented parsing* (U-DOP) (Bod 2006), in the sense that U-DOP can generate any tree-structure with no lexical or structural constraints. However, while U-DOP requires costly computations to estimate the most probable parse trees, our model provides the same flexibility implemented in a simpler way.

A feature that we believe should be added to the model in the future is the ability to establish schemas based on the similarity of the strings and structures that enter the decision function. This feature should reduce learning costs, which will be necessary for processing natural language corpora with a much higher diversity and complexity than the small artificial languages used here. The schematizing feature may also generate increasingly abstract schemas organized in a hierarchical network that can be studied and compared with conventional grammatical descriptions of a language. A possible future application of the model is thus to describe the emergence of lexico-grammatical categories through language learning,

References

Bod, R. (2006, July). An all-subtrees approach to unsupervised parsing. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics* (pp. 865-872). Association for Computational Linguistics.

Bybee, J. (1985). Morphology: A Study of the Relation between Meaning and Form. Amsterdam: John Benjamins.

Bybee, J. (2002). Sequentiality as the basis of constituent structure. Typological Studies in Language, 53, 109-134.

Croft, W. (2001). Radical construction grammar: Syntactic theory in typological perspective. Oxford University Press on Demand.

Christiansen, M. H., Dale, R. A., Ellefson, M. R., & Conway, C. M. (2002). The role of sequential learning in language evolution: Computational and experimental studies. In Simulating the evolution of language (pp. 165-187). Springer, London.

Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. Behavioral and Brain Sciences, 39.

Frank, S. L., Bod, R., & Christiansen, M. H. (2012). How hierarchical is language use?. Proc. R. Soc. B, 279(1747), 4522-4531.

Ghirlanda, S., Lind, J., & Enquist, M. (2017). Memory for stimulus sequences: a divide between humans and other animals?. Royal Society open science, 4(6), 161011.

Grant, D. S., & Roberts, W. A. (1976). Sources of retroactive inhibition in pigeon short-term memory. Journal of Experimental Psychology: Animal Behavior Processes, 2(1), 1.

McCauley, S. M., & Christiansen, M. H. (2019). Language learning as language use: A cross-linguistic model of child language development. Psychological review, 126(1), 1.

MacDonald, S. E. (1993). Delayed matching-to-successive-samples in pigeons: Short-term memory for item and order information. Animal Learning & Behavior, 21(1), 59-67.

Roberts, W. A. (2002). Are animals stuck in time?. *Psychological bulletin*, *128*(3), 473.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

ILLUSTRATING THE CREATIVE ASPECTS OF SOUND SYMBOLISM: IMPLICATIONS FOR THEORIES OF LANGUAGE EVOLUTION

KEITH JONES[1] and MARCUS PERLMAN[*1]

[*]Corresponding Author: m.perlman@bham.ac.uk
[1]English Language and Linguistics, University of Birmingham, United Kingdom

This study examines how sound symbolism might have helped people create the first spoken words. An influential hypothesis is that sound symbolism played a key role in helping listeners connect spoken forms and meanings in the absence of convention. Experiments in the dominant kiki-bouba paradigm typically focus on comprehension as discrete mappings between forms and referents, and neglect creative aspects of meaning-making. In contrast, we illustrate the creative aspects of sound symbolism with a drawn replication of Köhler's (1947) classic experiment. Participants drew *kikis, boubas, taketes,* and *malumas* in two different contexts: as a creature or as a company symbol. Findings show that size, in addition to shape, is at play: *kikis* and *taketes* are both spikier and smaller than *boubas* and *malumas*. The salience of shape varies between contexts. *Kikis* and *taketes* elicited spikier drawings as symbols than as creatures. Thus, drawings are modulated by context, not just in superficial details, but in the very feature of shape under investigation in standard experiments. Our approach highlights the construction of meaning in context, a creative process that would have been crucial to the formation of the first words.

1. Introduction

The evolution of language is, in its essence, a creative process. Consider, for example, the formation of the first spoken words. At some point in human history, our ancestors established the original symbolic vocalizations used to express the details of their thoughts and experiences. These "proto-words", in addition to being created for the first time, needed to be understood by a listener without any prior code for how to interpret their meaning. Thus, without convention to guide them, listeners needed to make sense of the vocalizations, constructing their meaning from available context (including gestures) and common ground. Indeed, philosophers have observed that understanding the meaning of a word in an entirely foreign language is, logically speaking, a remarkably difficult problem, as listeners are faced with a near-endless set of possibilities for its potential meaning (Quine, 1960).

As a solution to this challenge, some researchers have hypothesized that sound symbolism – a cross-modal resemblance between speech sounds and meanings – played a key role in the comprehension of the first words (Cuskley & Kirby, 2013; Ramachandran & Hubbard, 2001; Imai & Kita, 2014). Köhler's (1929, 1947) classic experiment on shape-sound symbolism serves as the most influential and highly scrutinized case of how this works (Lockwood & Dingemanse, 2015). When presented with the spiky and rounded shapes in Figure 1, and asked which to call *takete* and which *maluma*, participants overwhelmingly associated *takete* with the spiky shape and *maluma* with the rounded one. Similar results, revamped by Ramachandran and Hubbard (2001) with *kiki* and *bouba*, are widely documented (Styles & Gawne, 2017), found with participants across ages (Imai et al., 2015) and cultures (Bremner et al., 2013).

The evident universality of the kiki-bouba effect has led scholars to argue that such intuitive mappings between spoken form (e.g. rounded consonants and vowels) and visual meaning (rounded shape) would place "natural constraints on the ways in which sounds are mapped on to objects" (Ramachandran & Hubbard, 2001: 19). Thus, sound symbolism could help to narrow a listener's search through semantic space, and thereby provide a critical clue to a speaker's meaning. This function of sound symbolism could have been particularly advantageous to human ancestors in the process of establishing the first vocal symbols.

In this paper, we consider how sound symbolism might have functioned to bootstrap (Imai & Kita, 2014) comprehension of the first spoken words, helping people to bridge the gap between form and meaning (Perniss & Vigliocco, 2014). Yet, while our study contributes positively to the broad enterprise of sound symbolism (Lockwood & Dingemanse, 2015), the launching point for our experiment is a critical observation about much of this research, particularly the extensively documented kiki-bouba effect. Many experiments that investigate this phenomenon operate with fixed, abstract shapes as stimuli, in which the semantic contrast of angularity is built into the design (see Westbury et al., 2018 for a thorough methodological critique). Often, as in the classic experiment of Köhler (1929, 1947), as well as Ramachandran and Hubbard's (2001) frequently cited informal replication, this contrast is accentuated in a two-alternative forced choice paradigm. Stemming from this, research on sound symbolism, and the theories that arise from it, often treat the process of comprehension as a matter of "mapping" between speech stimuli and their targeted meanings, operationalized in phrases like "cross-modal correspondence" or "association". This approach tends to background the core creative processes involved in how people construct meaning from words in context, such as sensorimotor simulation (Bergen, 2012; Hostetter & Alibali, 2008; Perlman & Gibbs, 2013) and open-ended inference (Sperber & Wilson, 1986; Sulik, 2018).

1.1. Current study

In the current study, we illustrate the creative aspects of sound symbolism with a drawn replication of Köhler's classic shape-symbolism experiment. Recently, Davis et al. (2019) found that drawings elicited by nonce words reflect sound-symbolic connections to the forms of these words. Here, we report a simple demonstration of how providing participants the freedom to draw their interpretations of *taketes* and *malumas* (and *kikis* and *boubas*) can reveal the constructive elements of meaning-making that are obscured in the standard forced-choice matching paradigm. When the contrast of angularity is not built into the experimental design, and participants are free to imagine the meanings of the words, do they still focus distinctly on shape? Or do other aspects of meaning come to mind? For instance, they might also attend to size: studies show that some of the same features of speech sounds that convey shape are also associated with magnitude, e.g., high, front vowels (Thompson & Estes, 2011) and voiceless consonants (Vigliocco & Kita, 2006) with small.

Moreover, meaning-making does not happen in a vacuum, and the sound symbolic cues of pseudowords might vary in different contexts that draw attention to different aspects of meaning. To examine the effect of context, we asked participants to produce their drawings in one of two different scenarios – as a label for a creature or as a symbol for a company. We then analyzed both the spikiness and the size of the drawings. Our findings highlight the constructive, creative process of making meaning from sound symbolism, in context, which would have been critical in the formation of the first words.

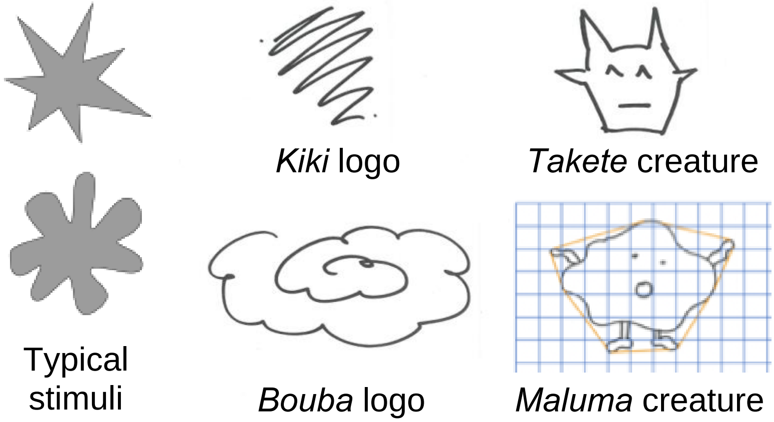


Figure 1. Left column shows typical experimental stimuli. Right columns show drawings from the current experiment. The bottom rightmost image shows how the drawing was enclosed to measure its spread.

2. Methods

2.1. Participants

52 participants were recruited on campus at the University of Birmingham in the United Kingdom. All were native English speakers. Three participants were later

excluded from analysis because they indicated previous familiarity with the kiki-bouba effect.

2.2. Materials and Design

The pseudoword stimuli used to elicit drawings included *kiki* and *bouba* from Ramachandran & Hubbard (2001) and *takete* and *maluma* from Köhler (1947). *Bouba* and *maluma* represented rounded words (consonants and vowels with lip rounding) and *kiki* and *takete* unrounded words (also characterized by voiceless plosive consonants). Each word was printed on a 4-by-6-inch white card. (Thus, notably, participants were presented with written words, a point to which we return in the Discussion.) Participants were randomly assigned to one or the other word set, and then to one of the two contextual scenarios. In the *symbol* scenario, it was explained that "An entrepreneur has founded two new brands," and in the *creature* scenario, that "A space explorer has discovered two new alien creatures." Participants were then asked to draw the two items based only on their names. The words were presented in counter-balanced order across participants. Participants drew on blank index cards fastened to a clipboard, using a black fine point marker.

2.3. Procedure

Participants were approached outdoors on the university campus, where they were asked to complete a brief experiment. After agreeing to participate, they signed a consent form and the experiment began. Participants first read their scenario, and then were handed the marker and clipboard on which to produce their drawings. They were then presented with the words, one at a time. Participants had 20 seconds to complete each drawing. At the end, participants provided written responses to three questions: (1) Can you describe both of your [symbols / creatures]?; (2) Can you explain how you decided on the size and shape of your [symbols / creatures]?; and, (3) Have you heard of the 'kiki-bouba effect'?

2.4. Analysis

We conducted quantitative analyses of the shape and size of the drawings. For shape, we measured their degree of angularity by soliciting judgments from three naïve raters. The raters viewed the drawings one at a time in randomized order. For each, they rated its shape on a scale from 1 to 7, with 1 being most spiky and 7 most round. Analyses used the average of the three ratings for each drawing.

To gauge the size of the drawings, we measured the spread of the surface area they covered. Using a photocopier, each drawing was transferred onto gridded paper (1 cm_2) and enlarged by 200%. The spread was measured by enclosing the drawing with a (hand-drawn) convex polygon that contacted its outward points (Figure 1). We then counted the number of grid-squares inside the enclosed area, with bisected squares counted when half or more was enclosed. The total spread of the drawing was the sum of the counted grid-squares.

Statistical analyses were conducted in R version 3.4.3 (R Core Team, 2015), and mixed effects models analyses were performed with the lme4 package version 1.1-21 (Bates, Maechler, Bolker, & Walker, 2015). Statistical significance was determined using model comparisons with and without the factor of interest.

3. Results

Examples of drawings for each word and scenario are shown in Figure 1.

3.1. Quantitative Analysis of Shape and Size

Figure 2 displays means and standard errors of the shape ratings and spread. These include results of each word separately and combined together, in both contextual scenarios. The complete statistical analysis and results can be found at https://osf.io/z7ty4/. Here, we report the overall results for the four words combined. To test whether word type affected the shape and size of the drawings, and whether this varied between contexts, we constructed two linear mixed effects models, one with shape ratings and one with spread as the dependent measure. In each, fixed effects included word type (rounded / unrounded), context (creature / symbol), and a factor for their interaction. Random intercepts were added for participant and word set (kiki-bouba / takete-maluma), as well as a random slope for word type by word set. Independent variables were centered.

The results show that drawings of unrounded words were rated as spikier than drawings of rounded words, $b_0 = -1.99$, *s.e.* = 0.32, $t = -6.31$, $p = 0.007$. This effect was modulated by context: unrounded words were rated as spikier in the symbol condition, $b_0 = -1.23$, *s.e.* = 0.58, $t = -2.14$, $p = 0.033$. There was a marginal main effect of context on the shape ratings, $b_0 = -0.55$, *s.e.* = 0.30, $t = -2.14$, $p = 0.068$.

For size, we found a main effect of word type on the spread of drawings: drawings for unrounded words were smaller than rounded words, $b_0 = -25.6$, *s.e.* = 7.4, $t = -3.46$, $p = 0.019$. There was no interaction between word type and context, and no main effect of context.

3.2. Participants' Descriptions

To gain further insight into the motivation for the drawings, we coded participants' responses to the post-experiment question asking them to describe their drawings. The adjectives they used were categorized as relating to spiky or curved shape and to small or large size. For example, spiky words like “angular”, “sharp”, “jagged”, “pointy” and “spiky” were among those adjectives used for *takete* symbols. *Bouba* creatures were described as “long”, “large”, “blob-like”, and “bulbous”. Table 1 presents a summary of these results.

Table 1. Frequency of shape and size words used in participants' descriptions.

Word	*Creature Shape (Spiky / Curved)*	*Symbol Shape (Spiky / Curved)*	*Creature Size (Small / Large)*	*Symbol Size (Small / Large)*
Kiki	0 / 1	11 / 0	5 / 1	0 / 0
Bouba	0 / 3	0 / 8	1 / 6	0 / 2
Takete	4 / 0	8 / 0	3 / 5	0 / 0
Maluma	0 / 7	0 / 9	0 / 2	0 / 0

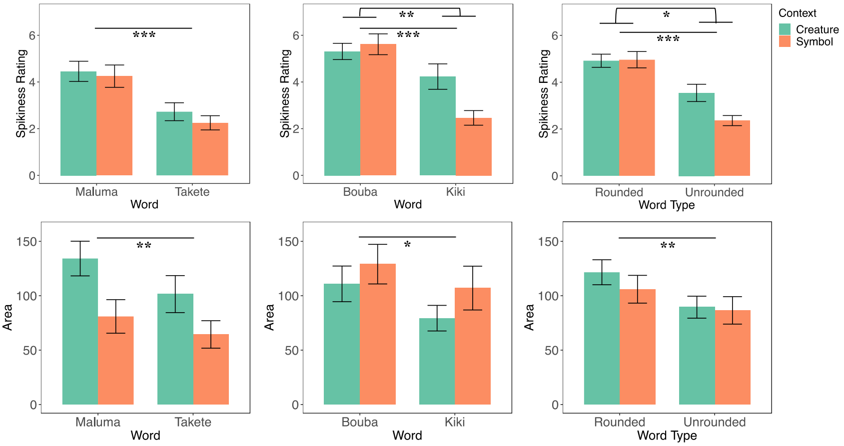


Figure 2. Means and standard errors of the shape ratings (top) and spread (bottom). Creatures are shown in green, symbols in orange. Lines show significant differences between the word types, and brackets show interactions between word and context. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

4. Discussion and Conclusion

Our study launches from the critical observation that the seminal paradigm of experiments on sound symbolism, the kiki-bouba effect, is based on a design that builds in discrete representations of meaning. This reduces the creative aspects of comprehension – for example, processes of meaning-making such as sensorimotor simulation or open-ended inference – to the discrimination of specific semantic contrasts. Thus, by its design, the paradigm reinforces the notion that the primary function of sound symbolism is to facilitate mapping between forms and referents, rather than focusing on sound symbolism as a guiding cue in the construction of meaning. To contrast these two perspectives, we examined whether, by giving participants more freedom to construct meaning from pseudowords, we could uncover aspects of sound symbolism that are lost in many standard experiments. Such an approach might provide fresh perspective into the formation of the first vocal symbols by shifting focus to more creative aspects of sound symbolism.

In our replication of Köhler’s classic experiment, participants drew pictures of *kikis, boubas, taketes,* and *malumas* – either as an alien creature or as a symbol for a company. Analysis of their drawings replicates the typical shape-sound symbolism effect, but also reveals patterns that extend beyond this standard finding. Quantitative results indicate that multiple semantic dimensions – size in addition to shape – are at play in the representation of the words. *Kikis* and *taketes* are both spikier and smaller than *boubas* and *malumas*. Critically, the salience of the shape dimension varies between the two contexts: unrounded words are depicted as spikier as a symbol than as a creature. Thus, the representations that people draw are modulated by context, not just with respect to superficial details, but to the very feature of shape that is under test in standard shape-sound symbolism experiments.

The effect of context on the salience of the shape and size dimensions was also evident in participants’ descriptions of their drawings, which referred to key differences between representations of the words in the different scenarios. References to size were far more common in participants’ explanations for creatures, whereas this dimension was scarcely mentioned with respect to symbols. Thus, size appears to be a more salient semantic feature of creatures than company logos. In contrast, descriptions of symbols focused much more on shape. Thus, we see how context can modulate sound symbolism, drawing attention to different semantic dimensions depending on their relevance.

Critical readers will note that while our study aims to examine sound symbolism, participants based their drawings on written prompts. Therefore, there is the potential that their drawings were influenced by the orthographic shapes of the words, a possibility that is amplified by research showing that, even with auditory stimuli, orthography can play a major role in the kiki-bouba effect (Cuskley et al., 2017). To investigate this possibility in our results, we coded participants’ explanations of their drawings for whether they directly mentioned the orthography and the sound of words as motivation. Of the 49 participants (who had not heard of the kiki-bouba effect), 32 suggested that word-sound alone influenced the appearance of their drawings, while only three suggested that orthography alone had an influence and four referred to word-sound in combination with orthography. These data suggest that participants generally had the intuition that they were attending mainly to the sound of the words. Whether this is true is to be determined in future work, in which we plan to compare our results to the vocal presentation of the words.

In this study, we have presented a critique of the theoretical bias that is introduced by an experimental design that operationalizes comprehension as mapping words to discrete meanings along narrow semantic dimensions. For the purpose of demonstration, we have focused on the classic kiki-bouba effect because of its widely cited influence in sound symbolism research (Lockwood &

Dingemanse, 2015). There is, of course, a great deal of varied work on sound symbolism, in methodology, as well as in the phenomena of study (Motamedi et al., 2019). Sound symbolism goes far beyond visual shapes, with experiments finding that people associate different phonemes with properties related to dimensions like brightness, taste, speed, and precision, to name a few (Lockwood & Dingemanse, 2015). In some experiments, participants are able to modulate the meaning space in their response, for example, adjusting the speed of a moving ball (Cuskley, 2013), and recent work has also elicited sound-symbolic drawings (Davis et al., 2019). Other experiments test large sets of word stimuli varying in phonetic properties (Ahlner & Zlatev, 2011), and present participants with large batteries of semantic dimensions (Monaghan & Fletcher, 2019; Klink, 2000; Westbury et al., 2018). Experiments also test the comprehensibility of real words in foreign languages, typically using translations of the words as alternative responses (Brown et al., 1956; Dingemanse et al., 2016). And still others investigate sound symbolism in the process of learning novel words (Nielsen & Rendall, 2012), including whether sound-symbolic labels help to categorize different stimuli (Lupyan & Casasanto, 2015). There is also considerable work examining sound symbolism across development (Kantarzis et al., 2011) and across cultures (Bremner et al., 2013). Yet, while this various research is highly informative, it still tends to operationalize comprehension of sound symbolism as a discrete mapping between the form of a word and its meaning or referent.

Our study builds on the premise that the evolution of language is, at its core, a creative process. Therefore, in order to understand the function of sound symbolism in language origins (e.g. the formation of words), we reason that it is important to examine its creative aspects. Indeed, the creative nature of sound symbolism is well documented outside of experiments on the kiki-bouba effect and the comprehension of sound symbolism. Studies of natural language use, particularly those in a multimodal framework, highlight the rich ways that sound symbolism manifests in human communication – although it is often discussed in somewhat different terms, e.g. demonstration, depiction, ideophones, mimicry, and iconicity (e.g. Clark, 2016; Dingemanse, 2013; Kendon, 2017; Laing, 2019; Lewis, 2009).

In this light, we hope that this simple demonstration – a drawn replication of Köhler’s classic experiment – has illustrated some of the creative aspects of sound symbolism that are obscured in the standard kiki-bouba paradigm. Critically, this approach, rather than formulating comprehension as a discrete mapping between forms and meanings, emphasizes the construction of meaning in context. We suggest that this kind of creativity might have been crucial to the formation of the first words, during a period when early humans could not rely on a well-established system of vocal symbols.

References

Ahlner, F., & Zlatev, J. (2010). Cross-modal iconicity: A cognitive semiotic approach to sound symbolism. *Sign Systems Studies, 38*(1/4), 298–348.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, (67), 1–48.

Bergen, B. K. (2012). *Louder Than Words: The New Science of how the Mind Makes Meaning.* Basic Books.

Bremner, A.J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition, 126*(2), 165–172.

Brown, R.W., Black, A.H., & Horowitz, A.E. (1955). Phonetic symbolism in natural languages. *Journal of Abnormal Psychology, 50*(3), 388–393.

Clark, H. H. (2016). Depicting as a method of communication. *Psychological Review, 123*(3), 324–347.

Cuskley, C., & Kirby, S. (2013). Synesthesia, Cross-Modality, and Language Evolution.

Cuskley, C., Simner, J., & Kirby, S. (2015). Phonological and orthographic influences in the bouba–kiki effect. *Psychological Research,* 1–12.

Davis, C.P., Morrow, H.M., & Lupyan, G. (2019). What does a horgous look like? Nonsense words elicit meaningful drawings. *Cognitive Science, 43*.

Dingemanse, M. (2013). Ideophones and gesture in everyday speech. Gesture, 13(2), 143–165.

Dingemanse, M., Schuerman, W., Reinisch, E., Tufvesson, S., & Mitterer, H. (2016). What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages. *Language, 92*(2), e117–e133.

Hostetter, A.B., & Alibali, M.W. (2008). Visible embodiment: Gestures as simulated action. Psychonomic Bulletin & Review, 15(3), 495–514.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal. Society B, 369*(1651), 20130298.

Imai, M., Miyazaki, M., Yeung, H.H., Hidaka, S., Kantartzis, K., Okada, H., & Kita, S. (2015). Sound Symbolism Facilitates Word Learning in 14-Month-Olds. *PLOS ONE, 10*(2), e0116494.

Kendon, A. (2017). Reflections on the “gesture-first” hypothesis of language origins. *Psychonomic Bulletin & Review, 24*(1), 163–170.

Klink, R.R. (2000). Creating Brand Names With Meaning: The Use of Sound Symbolism. *Marketing Letters, 11*(1), 5–20.

Kohler, W. (1970). *Gestalt Psychology: The Definitive Statement of the Gestalt Theory (2nd Revised edition edition).* New York: Liveright.

Lewis, J. (2009). As well as words: Congo Pygmy hunting, mimicry, and play. In R. P. Botha & C. Knight (Eds.), *The Cradle of Language*. Oxford University Press USA - OSO.

Lockwood, G., & Dingemanse, M. (2015). Iconicity in the lab: A review of behavioral, developmental, and neuroimaging research into sound-symbolism. *Frontiers in Psychology, 6*.

Lupyan, G., & Casasanto, D. (2015). Meaningless words promote meaningful categorization. *Language and Cognition, 7*(2), 167–193.

Monaghan, P., & Fletcher, M. (2019). Do sound symbolism effects for written words relate to individual phonemes or to phoneme features? *Language and Cognition*, 1–21.

Motamedi, Y., Little, H., Nielsen, A., & Sulik, J. (2019). The iconicity toolbox: Empirical approaches to measuring iconicity. *Language and Cognition*, 1–20.

Perlman, M., & Gibbs Jr, R.W. (2013). Sensorimotor simulation in speaking, gesturing, and understanding. In C. Müller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction* (Vol. 1, pp. 512–533). De Gruyter, Inc.

Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651).

Quine, W.V.O. (1960). *Word and object*. Cambridge, MA: MIT Press.

R Core Team. (2014). R: A Language and Environment for Statistical Computing. Retrieved from http://www.R-project.org/

Ramachandran, V.S., & Hubbard, E.M. (2001). Synaesthesia–a window into perception, thought and language. *Journal of Consciousness Studies, 8*(12), 3–34.

Sperber, D., & Wilson, D. (1986). Relevance: Communication and cognition (Vol. 142). Harvard University Press Cambridge, MA.

Styles, S.J., & Gawne, L. (2017). When Does Maluma/Takete Fail? Two Key Failures and a Meta-Analysis Suggest That Phonology and Phonotactics Matter. *I-Perception, 8*(4), 2041669517724807.

Sulik, J. (2018). Cognitive mechanisms for inferring the meaning of novel signals during symbolisation. *PLOS ONE, 13*(1), e0189540.

Thompson, P.D. & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. *Quarterly Journal of Experimental Psychology, 64*, 2392-2404.

Vigliocco, G. & Kita, S. Language-specific properties of the lexicon: Implications for learning and processing. *Language and Cognitive Processes, 21*, 790-816.

Westbury, C., Hollis, G., Sidhu, D.M., & Pexman, P.M. (2018). Weighing up the evidence for sound symbolism: Distributional properties predict cue strength. *Journal of Memory and Language, 99,* 122-150.

MODELLING CULTURAL EVOLUTION OF PRAGMATIC COMMUNICATION WHEN LANGUAGE CO-DEVELOPS WITH PERSPECTIVE-TAKING

ANNA KAPRON-KING*[1], SIMON KIRBY[2], and MARIEKE WOENSDREGT[3]

*Corresponding Author: annakk@cs.toronto.edu
[1]Computer Science, University of Toronto, Canada
[2]Centre for Language Evolution, University of Edinburgh, UK
[3]Centre for Language Studies, Radboud University, Nijmegen, Netherlands

Language use relies on the pragmatic ability to take into account an interlocutor's perspective when producing and interpreting utterances (Levinson, 1983). In natural language, semantic meaning can be underspecified (i.e. a given utterance can be compatible with multiple literal interpretations), and a specific interpretation is arrived at by means of pragmatic inference. Brochhagen et al. (2018) used a computational model of iterated learning to show that this division of labour can evolve when two different pressures are combined: a pressure for learnability and a pressure for communicative success. The pressure for learnability favours lexicons with simpler representations, and therefore makes lexicons with underspecified meanings more likely to evolve. The selection pressure for communicative success causes pragmatic communicators to be favoured, because these can compensate for the ambiguity of the signals in an underspecified lexicon by means of their pragmatic ability. In this model, both lexicons and communication types (i.e. literal or pragmatic) were culturally transmitted (through joint inference), and under the combination of the two pressures described above, they could co-evolve to produce the division of labour we see in natural language.

In a similar model, Woensdregt (2019) explored how lexicons and pragmatic ability evolve when accurate lexicon-learning depends on a co-developing ability to infer the speaker's perspective. In this model, learners were jointly inferring the lexicon and perspective of their cultural parent. Lexicons in this model were therefore culturally transmitted (as in Brochhagen et al., 2018), but pragmatic ability was transmitted genetically. Woensdregt showed that in this case, a similar division of labour to the one described above can arise both under a selection pressure for communicative success alone, and under a selection pressure for accurate perspective-inference alone.

Here, we demonstrate that language, pragmatic communication, and perspective inference can all simultaneously culturally evolve by combining features of both Brochhagen et al. (2018) and Woensdregt (2019). We are specifically interested in how lexicons and communication types evolve (through iterated learning) in the absence of any selection pressures. Following Woensdregt, our model treats communicative behaviour as the outcome of an interplay between the context in which communication occurs, the agent's individual perspective on the world, and the agent's lexicon. The combination of a given context and the speaker's perspective determines a probability distribution over potential referential intentions for the speaker. Each agent's perspective and lexicon are private mental representations, not directly observable by other agents. Language learners are therefore confronted with the task of jointly inferring both the lexicon and perspective of their cultural parent. Importantly, the learner always considers all referents as potentially being intended by the speaker, and can therefore not rely on cross-situational learning (Siskind, 1996) in order to infer the lexicon. Hence, the learner must rely on perspective-learning to learn the lexicon. Following both Brochhagen et al. and Woensdregt, we base our model of pragmatic communication on the Rational Speech Act model (Goodman & Frank, 2016), in which a speaker chooses their utterance by maximising the probability that the listener will interpret it as their intended referent. The addition of perspectives and contexts (following Woensdregt) means that pragmatic speakers choose their utterance not just based on the combination of their intended referent and their lexicon, but also on the context.

Two outcomes are plausible, which represent different divisions of labour between the culturally transmitted language and pragmatic inference by individuals in the population. We could have a lexicon of unambiguous one-to-one mappings being used by literal agents; or we could have a less-specified language being used by pragmatic agents. We show that, even in the absence of any selection pressures, the latter division of labour is a more likely outcome of cultural evolution. In other words, a language that relies on pragmatics evolves. Why is this? We argue that there are several possible lexicons that when combined with a literal speaker can lead to a learner inferring that the speaker is pragmatic and using a different lexicon. Importantly, the converse situation is less likely. More generally, this is because pragmatic communicators use their utterances more flexibly, depending not just on their intended referent and lexicon, but also on the context. Thus, once a pragmatic communication type has been adopted, it is unlikely to be confused with a literal communication type, because literal communicators use their utterances more strictly dependent on whether they are associated with the intended referent or not. Pragmatic communication is therefore an "attractor" in the space of culturally evolving systems (Sperber, 1996). Once it has evolved, populations may find it hard to retreat from it.

References

Brochhagen, T., Franke, M., & Rooij, R. van. (2018). Coevolution of Lexical Meaning and Pragmatic Use. *Cognitive Science*.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic Language Interpretation as Probabilistic Inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Levinson, S. C. (1983). *Pragmatics.* Cambridge University Press.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1–2), 39–91.

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*.

Woensdregt, M. S. (2019). *Co-evolution of language and mindreading: A computational exploration*. University of Edinburgh, Edinburgh.

MANIPULATING LEARNING CONSTRAINTS AND SIGNAL FEATURES IN ITERATED LANGUAGE LEARNING

VERA KEMPE [*1], SHEILA CUNNINGHAM [*1], NIKOLAY PANAYOTOV [*1],
NICOLAS GAUVRIT [*2] & MONICA TAMARIZ [3]

[*]Corresponding Author: v.kempe@abertay.ac.uk
[1]Division of Psychology, Abertay University, Dundee, Scotland, U.K.
[2]That Department, University Y, City, Country
[3]Department of Psychology, Heriot Watt University, Edinburgh, Scotland, U.K.

Iterated language learning (e.g. Kirby, Cornish & Smith, 2008) is a well-established methodology for studying emergence of linguistic structure in the laboratory. Yet there is considerable variation in implementation of the paradigm the effects of which have not been systematically scrutinized. For example, different training regimens can affect amount of learning thereby imposing different pressures on emergence of systematic signal-meaning associations. Also, differences in learning of visual vs. auditory signals (e.g. Raviv & Arnon, 2018) may not just be due to modality but to temporal differences affecting signal decay: the fading nature of sound may impose a memory burden not found in visual signals. We manipulated amount of learning, modality and temporal properties of the signals to examine the effects of these factors on the kinds of signal-meaning mappings that can emerge, while controlling for all other aspects of the methodology. We used binary sequences as signals to minimize potential influence of participants' prior knowledge and meta-linguistic awareness.

Both experiments alternated comprehension and production tasks to impose pressures to avoid ambiguity while also minimising effort: During comprehension, participants perceived the signals and had to choose the corresponding referents out of eight unfamiliar objects differing in size (large/small), shape (spiky/fluffy) and brightness (light/dark). During production, participants saw the referents and had to generate the corresponding signals, while being prevented from reusing the same signal for different referents. Crucially, amount of learning was manipulated by an adaptive procedure where transmission only took place once learners had reached a pre-defined accuracy threshold.

In Experiment 1, participants in 12 chains of 10 generations each learned to correctly associate either three (short training condition) or five (long training condition) meanings with signals comprising 4-6 bit binary auditory sequences of high and low tones of 500 ms each before being allowed to produce the signals that were then transmitted to the next generation. While compositional structure did not emerge, iconicity did: participants extended the length of those sequences that were associated with larger objects, but only in the short training condition.

In Experiment 2, participants in 12 chains of 10 generations each learned to associate 4-6-bit binary visual sequences comprising rows of blue and orange dots with the same meanings as in Experiment 1. In the stable condition, sequences appeared all at once before disappearing after a duration that was proportional to the number of dots in a sequence. In the fading condition, dots appeared one by one, with each dot fading away at a rate identical to the duration of tone presentation in Experiment 1. In all other respects, the procedure was similar to the short training condition in Experiment 1. As with auditory signals, there was a trend towards iconicity, but only in the stable condition. In the fading condition, participants tended to shorten all signals leading to overall greater learnability.

In sum, to simulate how memory constraints drive emergence of systematic mappings between signals and meanings in iterated language learning experiments training needs to aim at a ‘sweet spot’ where learning accuracy is high enough to retain crucial signal features yet also low enough to leave room for productive modifications. Our findings highlight the importance of carefully calibrating training regimens as well as the way in which modality-specific and temporal features can influence strategies of signal learning and signal use when trying to recreate language evolution in the laboratory.

Acknowledgements

This research was funded by Research Incentive Grant #70700 of the Carnegie Trust for the Universities of Scotland.

References

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, 105*(31), 10681-10686.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Developmental Science, 21*(4), e12593.

THE ROLE OF SIMPLICITY IN WORD ORDER HARMONY

DANIELLE NAEGELI [1*], RACHEL KINDELLAN [1*], KENNY SMITH [1], AND JENNIFER CULBERTSON[1]

[*]Joint First Authors, Corresponding Authors: *Danielle.Naegeli@ed.ac.uk*, *Rachel.Kindellan@ed.ac.uk*

[1]Centre for Language Evolution, University of Edinburgh, Edinburgh, UK

Previous research has demonstrated that the typologically attested preference for word order harmony, consistent order of dependents relative to their head, is reflected in individual learning behaviour in artificial language learning. To address the hypothesis that the word order harmony bias is driven by a more general bias for simplicity, we compared the strength of the word order harmony bias in one- and two-modifier contexts. It was predicted that the harmony bias would be stronger in the one-modifier conditions than in their two-modifier counterparts given that in one-modifier conditions the harmonic grammar is simpler than non-harmonic grammar, whereas in two-modifier phrases they are equally simple. While participants did not exhibit a uniform harmony bias in any condition based on frequency of majority order production, entropy calculations showed a difference in regularisation of harmonic and non-harmonic orders in the one-modifier conditions but not in the two-modifier conditions. The results presented here provide tentative evidence to support the hypothesis that a bias for word order harmony is rooted in simplicity.

1. Introduction

The preference for word order harmony–consistent order of heads and their dependents–has been well-attested cross-linguistically (Greenberg 1963, Dryer 1992). For example, languages with post-nominal adjectives also tend to have post-nominal numerals (Dryer, 2013a: Dryer, 2013b). Recent research has used artificial language experiments to study whether the typological preference for harmony in the noun phrase can be explained by a cognitive bias, e.g., for simplicity (Culbertson et al. 2012, Culbertson & Newport 2015, Culbertson & Kirby 2016); as a result of the transmission of languages through repeated generations of learners, languages then evolve to reflect these biases in learning.

Here we investigate the role of simplicity in driving the preference for harmony by focusing on a potential difference between grammars which generate noun phrases with a single modifier (e.g., an adjective *or* a numeral) and grammars which generate phrases containing multiple modifiers (e.g., an adjective *and* a numeral). Crucially, grammars producing harmonic one-modifier phrases can be shown to be representationally simpler than grammars producing non-harmonic counterparts. However, there is no clear difference in simplicity for harmonic and non-harmonic two-modifier grammars. If the harmony bias is rooted in simplicity, then a preference for word order harmony should be apparent when phrases contain only a single modifier, but weaker (or not present at all) when phrases contain two modifiers.

1.1. Previous experimental findings on noun phrase harmony

Several recent studies have conducted artificial language experiments to investigate whether learners prefer harmonic patterns of noun phrase word order. Culbertson & Newport (2015, 2017) showed that children taught a miniature artificial language comprised of phrases containing either a noun and an adjective *or* a noun and a numeral exhibited a strong preference for harmonic orders (i.e. where the adjective and numeral were placed consistently after the noun). In particular, they were better at learning harmonic input patterns, and shifted non-harmonic patterns to harmonic ones. Culbertson, Smolensky & Legendre (2012) showed that, albeit less strong, adults also have a similar harmony preference in this domain. Saldana, Smith, Kirby & Culbertson (2018) also found a preference for harmonic orders when participants were trained on noun phrases that had either one or two modifiers.

Culbertson & Kirby (2016) argued that the preference for harmonic orders may be rooted in a bias for patterns with simpler (more compressible) grammars. For example, a language which has post-nominal adjectives and post-nominal numerals can be captured by a general rule, e.g., NP → Noun Modifier (where → means 'expands to'). By contrast a non-harmonic language with post-nominal adjectives but pre-nominal numerals requires two specific rules, one for each modifier type, e.g., NP → Noun Adjective; NP → Numeral Noun. The harmonic grammar is shorter than the non-harmonic one, therefore it is simpler. This accords well with previous experimental results, where learners were trained (either exclusively or initially) with phrases containing only a single modifier.

However, there is no straightforward difference in simplicity for grammars producing two-modifier phrases. A learner must learn a rule for placement of the modifiers with respect to the noun as well as with respect to each other. Grammars producing two-modifier phrases following any order therefore consist of one specific rule (e.g., NP → N Adj Num, or NP → Num N Adj). If a cognitive preference for simplicity drives the typological prevalence of harmonic patterns, then this suggests a special role for single modifier phrases (cf. Hawkins 1994). In order to investigate this, we conducted an artificial language learning experiment where learners were exposed to either single modifier phrases or to two-modifier phrases. We predicted that a preference for harmonic orders would be stronger in participants in the one-modifier conditions than those in the two-modifier conditions.

2. Experimental test of the role of simplicity in harmony

2.1. Method

The experiment was closely modelled after Culbertson et al. (2012), however here we used a two-by-two design, manipulating whether the input language was predominantly harmonic or non-harmonic *and* whether participants were trained on one- or two-modifier phrases. The four conditions are illustrated in Table 1. We chose one harmonic pattern (both modifiers post-nominal) and one non-harmonic pattern (post-nominal adjective and pre-nominal numeral) in both one- and two-modifier contexts. As in Culbertson et al. (2012) each condition had a majority order used in the input. For the one-modifier conditions, the majority order for each modifier type was used 70% of the time (a direct replication of Culbertson et al. 2012). For the two-modifier conditions, there is single majority order, used 50% of the time, with all other possible orders used 10% of the time each.[1]

Table 1. Word order frequencies in input languages in each of the four experimental conditions shown by column. The majority word orders for each condition are in bold. In this two-by-two design, there are harmonic and non-harmonic majority orders in one- and two-modifier languages.

One-Modifier Conditions			*Two-Modifier Conditions*		
	Harmonic	*Non-Harmonic*		*Harmonic*	*Non-Harmonic*
Num **N**	0.3	**0.7**	**N** Adj Num	**0.5**	0.1
Adj **N**	0.3	0.3	**N** Num Adj	0.1	0.1
N Num	**0.7**	0.3	Num Adj **N**	0.1	0.1
N Adj	**0.7**	**0.7**	Adj Num **N**	0.1	0.1
			Num **N** Adj	0.1	**0.5**
			Adj **N** Num	0.1	0.1

Stimuli. The lexicon used in this experiment, a subset of that used in Culbertson et al. (2012), is comprised of four nonsense nouns corresponding to novel objects lacking an English label (“blifona”, “grifta”, “nerka”, “wapoga”), three nonsense adjectives (“cherg”,”geej”, “fush”) referring to furry, green and blue, and three

---

[1] To justify these input frequencies, consider the likelihood that a participant in the one-modifier harmonic condition would see a post-nominal adjective *and* a post-nominal numeral: NMod (0.7) x NMod (0.7) is equal to NModMod (~0.5).

nonsense numerals (“derf”,”kez”,”glawb”) referring to two, three and four. Visual stimuli were identical to those used in Culbertson et al. (2012), extended to two-modifier descriptions as illustrated in Figure 1.

Procedure. Participants were instructed that they would be learning part of a new language. Then they were trained on vocabulary and phrases in the language, as described below. The input for each participant was determined by the condition, explained in detail below.

Phase 1: Vocabulary Training. First participants were trained on nouns in isolation. On each trial a single grayscale noun was presented along with auditory and orthographic labels (40 trials total, 10 for each noun, random order, see Figure 1A). This was followed by a testing phase in which participants were asked to verbally produce a label for a noun presented on the screen (40 trials total, 10 for each noun, random order). Then participants were trained and tested in the same manner for the modifiers in isolation (60 trials total per phase, 10 for each modifier, random order, see Figure 1B, C).

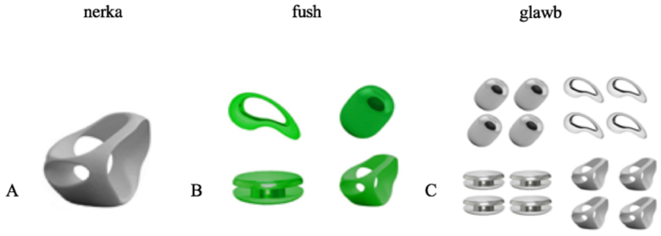


Figure 1: Example vocabulary training images for nouns (A), adjectives (B) and numerals (C). For adjectives, all 4 nouns were presented on screen modified by the relevant property. For numerals set of all 4 nouns were presented on screen in the relevant numerosity.

Phase 2: Noun & Modifier Training. Participants were then trained on phrases in the language. On each trial, a noun modified by either an adjective or a numeral (one-modifier condition) or both an adjective and a numeral (two-modifier condition) was presented on screen and the corresponding phrase was presented auditorily and orthographically. Figure 2 shows an example of each possible trial type.

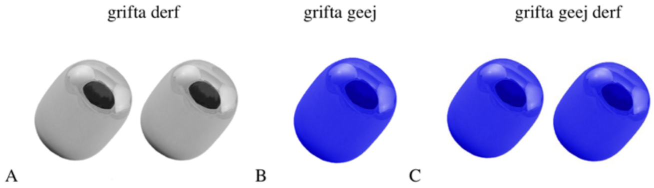


Figure 2: Noun and modifier phrase training examples. For the one-modifier condition numeral (A) and adjective (B) training. For the two-modifier condition numeral plus adjective training (C).

Phase 3 and 4: Picture Matching and Phrase Production. Participants were then tested on their knowledge of phrases in the language. This was done by alternating between comprehension and production. On each comprehension trial a description was presented along with four images: one target image depicting the target noun along with three foils, which varied in one modifier and noun. Participants were then asked to click on the image corresponding to the label (80 trials total, separated into four blocks of 20 trials). On each production trial an image appeared without a label and participants were asked to verbally describe it using the artificial language (80 trials total, separated into four blocks of 20 trials).

3. Results

3.1. Analysis of Majority Order Production

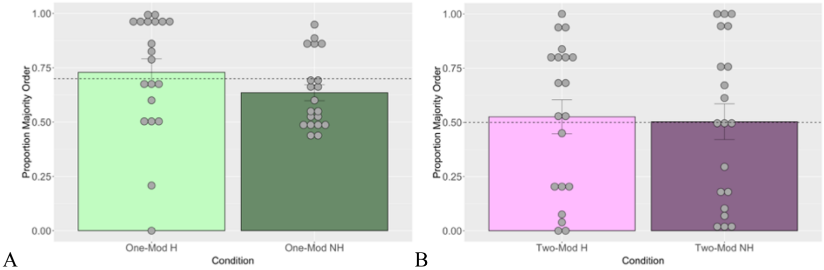


Figure 3. Proportions of majority order use in each condition are represented above. A and B represent one-modifier conditions and two-modifier conditions, respectively. Harmonic (H) and non-harmonic (NH) conditions are depicted on the left and right within the plots, respectively. The dotted lines represent the probability of the majority input order, 0.7 in (A) and 0.5 in (B). Points represent individual points and serve to demonstrate the variability across participants. Standard error is represented by the error bar.

Figure 3 shows how often participants in each condition used the majority order on test trials relative to the input frequency. A binary logistic regression model predicting use of majority order by harmony and number of modifiers (with by-participant random effects) revealed a significant effect of number of modifiers ($\beta = 0.517 \pm 0.252$, $p = 0.040$), simply reflecting the difference in majority order in the input, but no significant effect of harmony ($\beta = 0.161 \pm 0.252$, $p = 0.523$), and no interaction between harmony and the number of modifiers ($\beta = 0.270 \pm 0.$ 0.252, $p = 0.284$, respectively). This result is not in line with our predictions as it does not indicate a bias for harmonic word order in the one-modifier condition. However, a closer examination of individual data points reveals that behaviour across conditions was not uniform, and in fact has in some cases bimodal distribution. Figure 4 A and B show the different distributions of outcomes across harmonic and non-harmonic conditions for participants trained on one-modifier phrases; these results are reminiscent of Culbertson et al.'s (2012) results, in that participants in the harmonic condition tend to regularise using the majority order whereas very few participants in the non-harmonic condition regularise using the majority order. Figure 4 C and D show the different distributions of outcomes across harmonic and non-harmonic conditions for participants trained on two-modifier phrases. Like the one-modifier conditions, it appears that regularisation using the majority order is stronger in the harmonic than the non-harmonic order, though the difference in dispersion is less defined than between the one-modifier conditions. To explore these individual-level effects, a second analysis was adopted.

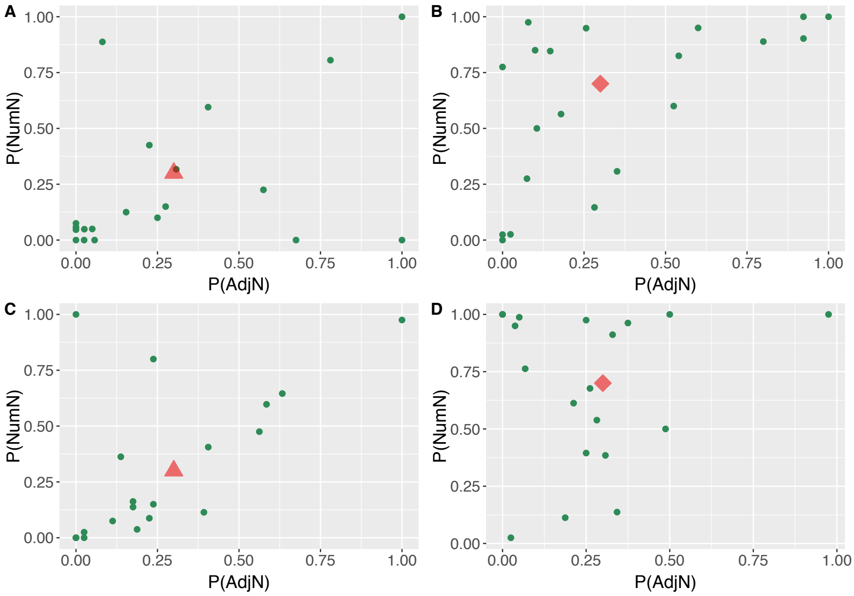


Figure 4. Individual participant outcomes as represented by probability of producing NumN (y-axis) and AdjN (x-axis) in the harmonic (A, C) and non-harmonic (B, D) conditions, for one-modifier (A, B) and two-modifier (C, D) languages. The input word order frequencies are represented by the triangle in (A, C) and by the diamond in (B, D).

3.2. Entropy Analysis

Shannon entropy provides a standardised measure of variation in a system (Shannon, 1948), calculated by taking the sum across all variants in a dataset of the probability of that variant times the log of that probability. We used this measure to assess change in variation in the use of the majority order from the input to the output language. Entropy of the input languages was first calculated. For the one-modifier languages the majority order has probability 0.7 and the minority order 0.3, yielding entropy 0.88. For the two-modifier languages the majority order has probability 0.5 and all minority orders combined have probability 0.5, yielding an entropy of 1. Then, entropy of each participants' output was calculated by collapsing all word orders into either majority order or minority order. This measure allowed for examination of change in variation by use of the majority order. Finally, change in entropy from input to output for each participant was calculated, shown in Figure 5. Unlike the analysis of majority order production reported above, this change in entropy calculation captures the fact that participants in the non-harmonic one-modifier condition maintain variation more and regularise using the majority order less than the one-modifier harmonic counterpart. By contrast, most of the participants in the one-modifier harmonic condition regularise on the majority input order, therefore reducing variation and entropy. This difference between harmonic and non-harmonic orders in the one-modifier conditions is as predicted; it shows a greater reduction in variation by overuse of the harmonic order. Also in line with our hypothesis, no such difference between harmonic and non-harmonic languages is observed in the two-modifier conditions.

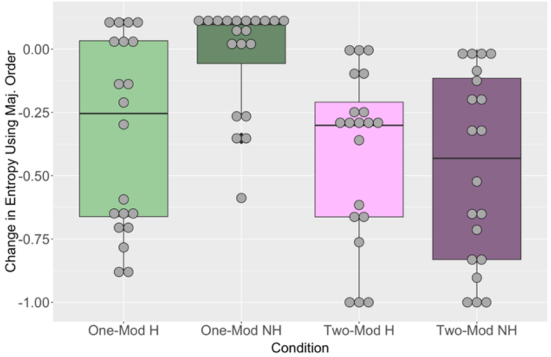


Figure 5. Change in entropy (y-axis) calculated using majority order represented per condition (x-axis). Points represent individual participants. Mean is represented by the bold horizontal line.

Change in entropy across conditions was analysed using a linear regression model. This model confirmed a significant effect of harmony ($\beta = 0.314 \pm 0.106$, $p < 0.01$) and an interaction between harmony and number of modifiers, such that the effect of harmony is stronger in the one modifier conditions than in the two modifier conditions ($\beta = -0.373 \pm 0.149$, $p = 0.015$)[2]. In other words, as predicted, a harmony preference was detected in the one-modifier conditions but not the two modifier conditions.

---

[2] Entropy Change ~ Harmony * Number of Modifiers

4. Discussion

Here we have reported the results of an artificial language learning experiment testing the hypothesis that a cognitive bias for harmony is driven by simplicity. This hypothesis predicts that a bias for harmony should be found (more strongly) across single-modifier phrases, since a grammar encoding harmonic order in such phrases is representationally simpler than one encoding non-harmonic order. By contrast, a grammar with harmonic order in phrases with multiple modifiers is not straightforwardly simpler to represent than one with non-harmonic order. Following previous work (Culbertson et al. 2012), participants were either trained on a predominantly harmonic language or a non-harmonic one. In addition, half of our participants were trained on one-modifier phrases only, and the other half on two-modifier phrases only. An analysis of majority order use failed to reveal the predicted difference between one- and two-modifier conditions in the preference for harmony. However, this analysis is somewhat misleading given that individual participants behaved quite differently from one another (e.g., see also Culbertson & Newport 2015). To deal with this variation, we turned to Shannon entropy to measure the extent to which participants across conditions reduced the variability in word order (comparing the use of majority input order vs other orders). This analysis revealed the predicted effects: significantly more reduction of entropy using majority pattern in the harmonic one-modifier condition compared to the non-harmonic one-modifier condition, but no difference in the two-modifier conditions.

Interestingly, the overall effect of harmony found in our one-modifier conditions appears to be weaker than the effect reported in Culbertson et al. (2012), who found a difference between harmonic and non-harmonic conditions even in the analysis of majority order production. Our guess is that this reflects a difference in the difficulty of acquiring the systems; the size of the lexicons used here was much smaller (4 nouns and six modifiers) compared to Culbertson et al (2012) (10 nouns and 10 modifiers). This may lead to a reduction in the overall effect of simplicity of the grammar, as participants spend fewer cognitive resources on vocabulary learning. Previous studies have also shown that greater cognitive load may lead to stronger regularisation behaviour (Hudson Kam & Newport, 2005; Ferdinand et al., 2019).

To conclude, while analysis of majority order proportions did not reveal a strong effect of harmony, the entropy analysis indicated that the harmonic order is preferred in one-modifier but not two-modifier contexts. This supports the hypothesis that a preference for harmony is rooted in a more general cognitive preference for simplicity, a well-known and wide-spread bias (Chater and Vinyati, 2003; Culbertson and Kirby, 2016). Given that the vast majority of modified noun phrases learners hear will include only a single modifier (Culbertson et al. under review), these phrases therefore have the potential to shape the evolution of noun phrase word order.

References

Chater, N., & Vitanyi, P. (2003). Simplicity: A unifying principle in cognitive science. Trends in Cognitive Science, 7, 19–22.

Culbertson, J. & Newport, E. (2015). Harmonic biases in child learners: In support of language universals. *Cognition, 139*, 71-82.

Culbertson, J. & Newport, E. (2017). Innovation of Word Order Harmony Across Development. *OpenMind, 1*(2) 91-100.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*(3), 306-329.

Culbertson, J., & Kirby, S. (2016). Simplicity and Specificity in Language: Domain-General Biases Have Domain-Specific Effects. *Frontiers in psychology*,*6*, 1964.

Culbertson, Jennifer, Marieke Schouwstra, and Simon Kirby. (2019). From the World to Word Order: The Link Between Conceptual Structure and Language. *PsyArXiv*. doi:10.31234/osf.io/v7be4.

Dryer, M. (1992). The Greenbergian word order correlations. *Language*, 68(1):81–183.

Dryer, M. (2013a). Order of adjective and noun. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved June, 2019 from http://wals.info/chapter/87.

Dryer, M. (2013b). Order of numeral and noun. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. Retrieved June, 2019 from http://wals.info/chapter/89.

Ferdinand, V., Kirby, S., & Smith, K. (2019). The cognitive roots of regularization in language. Cognition, 184, 53-68.

Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. Greenberg, editor, *Universals of Language*, pages 73–113. MIT Press, Cambridge, MA.

Hudson Kam, C., & Newport, E. (2005). Regularizing unpredictable variation. Language Learning and Development, 1, 151–195.

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? Journal of Experimental Psychology, 35, 815–821.

Saldana, C., Smith, K., Kirby, S. & Culbertson, J. (2018). Is regularisation uniform across linguistic levels? Comparing learning and production of unconditioned probabilistic variation in morphology and word order. Manuscript submitted for publication.

THE LEXICON AS A HUFFMAN CODE: WORDS ARE STRUCTURED FOR PROBABILISTICALLY BALANCED CONTRASTS

Adam King*[1] and Andrew Wedel[1]

*Corresponding Author: adamking@email.arizona.edu
[1]Department of Linguistics, University of Arizona, USA

1. Background

The key requirement for an efficient communication system is the maximization of communicated information across the message, relative to its length (Shannon, 1948). For simplicity, words can be thought of as messages, meaning that for efficient linguistic communication, word-internal information should be densely encoded within sub-word units.

As it happens, the lexicons (roughly, the set of words in the language) of nearly all tested languages display a robust relationship between word probability and length, i.e., *Zipf's law of abbreviation* (Zipf, 1935; Piantadosi, Tily, & Gibson, 2011; Bentz & Cancho, 2016), suggesting that the pressure for efficient communication partially affects the shape of words, at least with respect to the number of sub-word units. However, while short words may be part of an efficient system, they do not entail efficiency per se. For example, a language where every word is /ba/ satisfies the pressure for short forms, but fails *a priori* as a communication system. To this end, if the lexicon is structured for efficient communication, the sub-word units that build up full word forms should also be structured to be maximally informative, in addition to a general pressure for short lengths.

By definition, the average information for part of a message (in this case, contrastive sub-word units) is maximal when each possible contrast is equiprobable, such as in a Huffman code (MacKay, 2003). Therefore, if the lexicon is structured to be an efficient communicative code, contrasts within the lexicon's words should be between a more or less probabilistically balanced set of phonological contrasts.

2. Methodology

In this talk, we will show that contrasts in the lexicon significantly demonstrate effects of balance between contrastive phonological material, using a dataset of phonetically transcribed lexicons of 25 typologically diverse languages. Furthermore, we will demonstrate that the predicted trend towards balance extends beyond what might be expected in any language-like code, by comparing to a baseline for each language (Prado Martin, 2013).

Primarily, we will show an inverse relationship between type and token frequencies for 95% of the most probable contrasts[1], a relationship expected of a probabilistically balanced distribution, and importantly, one that does not require a prior specification of "*how balanced*" it need be. Assuming that the probability for each contrast is more or less equal, those that are associated with fewer word types should be associated words that have a higher frequency, on average (Fig. 1).

As an example, consider the segment [ð] in English which begins few words but, of those it begins, many words are high frequency function words, e.g., *this*, *the*, etc. [t], on the other hand, begins many more word types, though the average frequency of [t]-initial words is less. When put together, the overall probability of a word token beginning with [ð] or [t] is more or less balanced, at least more so than if the relation between type and token frequencies was different.

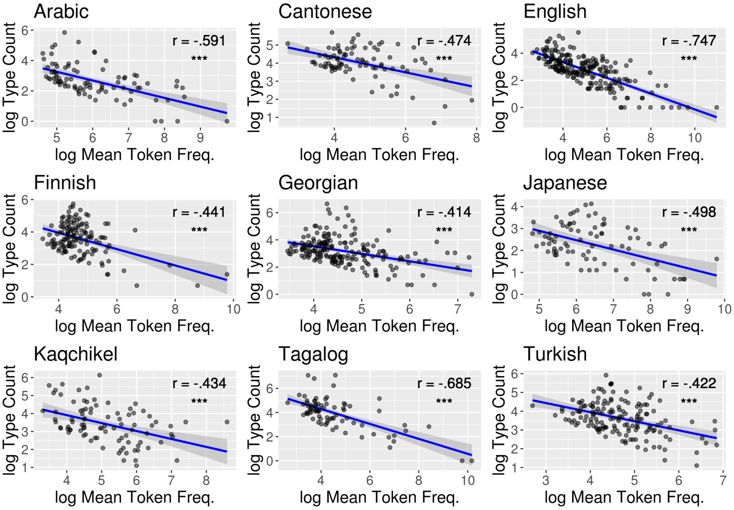


Figure 1. Inverse relation between type count and average frequency for word-initial biphone (2 segment sequences) in a subset of the 25 tested languages.

---

[1] This avoids effects of the so-called *Zipfian tail*, which creates an uninteresting positive relationship between type and token frequencies.

References

Bentz, C., & Cancho, R. Ferrer-i. (2016). Zipf's law of abbreviation as a language universal. *Proceedings of the Leiden Workshop on Capturing Phylogenetic Algorithms for Linguistics.*

MacKay, D. J. (2003). *Information theory, inference and learning algorithms.* Cambridge university press.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108*(9), 3526–3529.

Prado Martin, F. Moscoso del. (2013). The missing baselines in arguments for the optimal efficiency of languages. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35.*

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379 - 423.

Zipf, G. K. (1935). *The psycho-biology of language.* Mifflin Houghton Publishing.

CUMULATIVE CULTURAL EVOLUTION OF SYSTEMATIC STRUCTURE AND GRAMMATICAL COMPLEXITY IN HUMANS AND BABOONS

SIMON KIRBY*[1], VANESSA FERDINAND[2], KENNY SMITH[1], JOËL FAGOT[3] and NICOLAS CLAIDIÈRE[3]

*Corresponding Author: simon.kirby@ed.ac.uk
[1] Centre for Language Evolution, University of Edinburgh, UK; [2] Psychological Sciences, University of Melbourne, Australia; [3]Aix Marseille University, CNRS, France

Human language exhibits widespread systematicity at all levels of analysis. In phonology, syntax, and semantics, we see constraints on the allowable forms in the language such that the existence of some expression in the language is dependent on other expressions in that language. For example, the past tense ending “-ed” in English on a particular verb is used because of its appearance on other verbs in the language. This systematicity, a defining feature of language but rare in the behavioural repertoires of other animals, is a major focus of research in the field of language evolution. One approach shows that cultural evolution favours systematic structure when sets of behaviours are transmitted by iterated learning (e.g., Kirby et al., 2015). Cornish et al. (2013) explore the origins of systematicity in sets of sequences transmitted by iterated learning. Participants (adult humans) were shown and immediately attempted to recall a sequence of coloured lights. They are exposed to 60 such sequences and given feedback after each attempted recall. For the first participant in each chain of transmission, the 60 sequences were completely independent and random, but subsequent participants were given the sequences of the previous participant to copy. Although the sequences in the set are initially independent of each other, after several generations of transmission, they have become systematically structured, with commonalities across the strings making them easier to copy. As a result, the set of strings becomes increasingly *compressible* over generations.

Results like this have led researchers to propose that a universal learning bias for simplicity, driving cultural evolution through iterated learning, can explain a wide range of linguistic phenomena (see, e.g. Culbertson & Kirby, 2016). In this view, systematic languages have lower *grammatical complexity* than unsystematic ones: systematicity allows for a more concise set of rules to generate a language. However, we reanalyse the data from Cornish et al. (2013) to show that the increasingly compressible sets of sequences that emerge are actually underpinned by *more* complex grammars. We estimate grammatical complexity by using hidden markov model (HMM) induction (DeDeo, 2016) and find a gradual, cumulative increase in the complexity over generations (figure 1). Culture leads to increasingly compressible languages, but achieves this by introducing cumulatively more complex underlying structure.

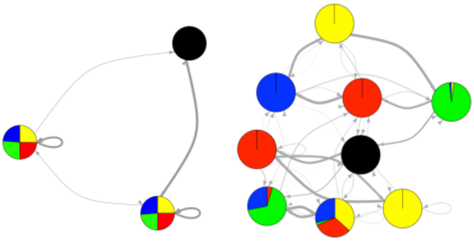


Figure 1. A typical HMM induced from the start of the human experiment (left) and one after 10 generations of cultural transmission (right). The probabilities of each of the colours being emitted by the hidden state are shown as pie charts, and the probabilities of transition from state to state by thickness of lines. The black node corresponds to the sequence separator. Over generations, complexity as measured by the number of hidden states increases.

To see if this process of cumulative evolution of systematicity underpinned by increasing grammatical complexity is unique to humans or a result of more widespread cognitive processes, we replicated the iterated sequence learning experiment in a population of captive baboons (Fagot & Bonté 2010), using a minimally adapted version of the paradigm from Cornish et al (2013). As in the human experiment, systematic structure in sequences emerged cumulatively over generations, resulting in compressible sets of behaviours. Once again, this is not the result of the underlying grammars becoming simpler over generations. However, unlike in the human data, there is no evidence for a cumulative increase in grammatical complexity, indicating a difference between humans and baboons in the way sets of structured behaviours culturally evolve.

These results add nuance to our understanding of what cultural evolution does to sets of behaviours as they are passed on by iterated learning. They support our conclusion that cultural evolution leads to increasingly compressible behaviour over generations, explaining the origins of systematicity in language. However, this does not necessarily mean that grammars will always become simpler, particularly in the case where the starting point is highly stochastic. We suggest that in some cases grammatical complexity might be an adaptation by cultural evolution to create rich, expressive languages that are nevertheless predictable. Understanding precisely what the situations are where grammatical simplicity or complexity is expected is an important target for future experimental research.

References

Cornish, H., Smith, K., & Kirby, S. (2013). Systems from sequences: An iterated learning account of the emergence of systematic structure in a non-linguistic task. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, No. 35).

Culbertson, J., & Kirby, S. (2016). Simplicity and specificity in language: Domain-general biases have domain-specific effects. *Frontiers in Psychology*, 6, 1964.

DeDeo, S. (2016). Conflict and computation on Wikipedia: A finite-state machine analysis of editor interactions. *Future Internet*, 8(3), 31.

Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (Papio papio). *Behavior research methods*, 42(2), 507-516.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, 141, 87-102.

THE ROLE OF RITUAL IN THE EVOLUTION OF LANGUAGE AND THE EMERGENCE OF RECURSION

ROBERT KLUENDER*[1] and EMILY DAVIS[1,2]

*Corresponding Author: rkluender@ucsd.edu
[1]Department of Linguistics, UC San Diego, La Jolla, CA, USA
[2]Centre for Language Evolution, University of Edinburgh, Edinburgh, Scotland

Human language presents a unique conjunction of referential and structural properties, the combined evolution of which is hard to account for. Here we synthesize older linguistic theories on the ritual origins of language with newer experimental results to present an exaptive account of both referential and structural aspects of language (particularly center-embedding) in ritual contexts. We discuss our current ongoing work using iterated learning tasks to elucidate the conditions under which novel center-embedded structures emerge, and conclude with a more general proposal for closer coordination of the language evolution research interests of anthropologists, ethologists, biologists and linguists in the study of ritual.

1. Introduction

One of the enduring conundrums for language evolution proposals is the thorny issue of how to account for the emergence of complexity on multiple, distinct yet interconnected levels simultaneously – or for that matter, even sequentially. Our goal in this paper is to highlight, revive and argue for the viability of language evolution proposals offered by linguists in the 1980s based on properties of animal and human ritual. We rely on commonalities outlined in Tonna et al. (2019) to generalize across animal and human ritual in this way: repetitive, sequential action chains in response to unpredictability, with emphasis on formal structure, precedence of exactitude of performance over any originally intended functional goal, and promotion of communication/bonding. Our main argument is that linguistic ritual-based proposals such as these offer plausible scenarios for the simultaneous development of both the crucial referential and structural properties of human language, especially recursion. We further investigate the claim that center-embedding is ubiquitous in human and animal ritual, as well as within the cognitive capacity of non-human species, and describe our progress thus far in trying to tease out experimentally the conditions under which center-embedding in particular might emerge. Finally, we highlight the prospects for more closely coordinating the study of the role of ritual in the evolution of language among social and cultural anthropologists, ethologists, biologists and linguists alike.

2. Brandon & Hornstein (1986)

Brandon & Hornstein (1986) proposed a three-stage scenario for the emergence of arbitrary symbolic reference from iconic systems of representation, via a process of ritualization in non-human animal species. The first stage occurs on the part of the receiver: an organism begins to treat the functional action of another organism, or a part thereof, as a sign of the action itself. For example, the functional movements necessary for initiation of bird flight (crouching, raising the tail, spreading the wings) can be interpreted causally by one bird as an indication that another bird intends to lift off. The second stage involves the sender accentuating and stylizing a behavior beyond its basic functional requirements in order to signal intention to a conspecific, which then interprets it as such. For example, Lorenz (1977:211) reports that pigeons exaggerate preflight behaviors in order to facilitate coordination of flight behavior with conspecifics within the same flock. In the third stage, a stylized behavior is displaced (in the ethological rather than the referential sense) from its source domain into a new sphere of activity. Bird courtship displays frequently incorporate stylized feeding and/or nesting behaviors, and primate displays that regulate social relations within a hierarchical dominance structure often incorporate stylized sexual behaviors.

Tomasello (2008:23) proposed ontogenetic versions of the same three-stage ritualization process. A displaced behavior can thus begin to look arbitrary to the outside observer – and possibly to the individual animal – as the original connection to the source domain or context is lost over time. If the connection can only be recovered through historical reconstruction, then for all practical purposes it becomes arbitrary in nature. Brandon & Hornstein (1986) speculate that this is a pre-adaptation for the emergence of purely symbolic representation.

3. Staal (1979, 1980, 1984a, 1989)

In a series of papers appearing in relatively obscure journals and anthologies (Staal, 1979, 1980, 1984a), and in a subsequent book (Staal, 1989), Frits Staal, a philosopher of language, linguist and Sanskritist, made a number of revolutionary claims about the evolution of language, based in particular on his study of Indian ritual. His views were informed by his observation and extensive documentation of a rarely performed twelve-day Vedic rite undertaken in 1975 in Kerala, a state in the deep southwest of India, where the ancient Vedic tradition that predated

Hinduism has been more prodigiously preserved by the priestly class than elsewhere on the subcontinent.

Staal's basic thesis is that ritual and the behaviors (and in the case of human ritual, even the language) associated with it are essentially meaningless. Like Brandon & Hornstein (1986), Staal (1979:12-13) argues that subsequent to displacement of particular behaviors from a source to a target domain in animal ritual, they lose their functional significance. While there may still be an expected outcome involved in rituals, even human ritualistic activity persists beyond the point at which the original motivation for it has become lost or is no longer valid (e.g. "eternal" flames). Staal (1979, 1984a) further argues that while ordinary activity focuses on results, ritualistic activity is focused solely on the correctness of performance. Anyone who has seen nature films of the courtship displays of bowerbirds can attest to the fact that there is extreme pressure to perform to acceptable standards – with severe selective consequences if those targets are not met – even in animal ritual.

Staal (1979) therefore concludes that the hallmark of ritual is its rule-governed nature. This leads to his further conclusion that it is the *structure* of ritual that matters far more than its purported content. His analysis of Vedic ritual demonstrates that it exhibits properties of both complementation, in the form of embedded sequences, and modification, i.e. sequences that undergo various changes in different contexts. In other words, it has the properties of a syntactic system. Staal goes on to hypothesize that the syntax of language was exapted from the syntax of ritual. Note that this proposal is not all that different in kind from that of Hauser, Chomsky, & Fitch (2002:1578) who speculate that the recursive properties of language could have been exapted from domains of animal behavior other than communication, such as navigation, numerosity, or social cognition.

However, Staal (1980) points out that ritual exhibits only center-embedding rather than right- or left-branching recursive structures. He suggests that since ritual unfolds on a much longer time scale than human language, there are no intrinsic working memory constraints on the use of center-embedding in ritualistic contexts, which seems plausible. He also suggests that the limited occurrence of center-embedding in human language may be a remnant of its unrestricted use in ritual. We suggest here that Staal might have actually taken this argument further.

The problem with the evolution of syntax has always been what the seed could be that would precipitate the emergence of structure in a relatively unstructured referential system. Staal's account already provides this seed via exaptation from ritual contexts. But once center-embedding has been established as a structural principle, it seems to us that its evolutionary transmogrification into right- or left-branching structures should be a fairly trivial matter. The virtually impossible center-embedded sequence (1) can easily be rendered transparent as (2).

(1) The woman [the man [the host knew __] brought __] left.
(2) The host knew the man [who brought the woman [who left]].

Yet this seems to be more of a referential (i.e. similarity-based interference) than a syntactic problem anyway, since changing the nature of the discourse referents renders even the use of center-embedded syntax transparent (Bever, 1970):

(3) The woman [someone [I knew __] brought __] left.

Note that Brandon & Hornstein's (1986) account of how symbolic reference could have emerged as a result of the ethological displacement of animal behavior from one domain to another in ritualized contexts provides an avenue for the referential properties of language to develop simultaneously and in parallel with syntactic properties of recursion via one general process of exaptation. To our knowledge, this is the only exaptive scenario that has this two-for-the-price-of-one advantage.

The next obvious question, however, is whether there is adequate evidence for center-embedding not only in forms of everyday human ritual less elaborated than complex Vedic rites, but also in animal behavior. We turn to this question next.

4. The evidence for center-embedding in human and animal behavior

4.1. Human ritual

Center-embedded sequences are more common in contemporary, everyday human ritual activity than may be apparent at first blush. Pulvermüller (2014) demonstrates how something as trivial as the quotidian ritual of toothbrushing can plausibly be analyzed as having up to ten layers of center-embedded action sequences, each consisting of matching starting and ending actions: e.g. taking out a toothbrush and putting it back into its holder, or opening and closing a door. Many formalized human ritual activities also exhibit this so-called ABA structure, in which a sequence is bracketed off by matching components at its beginning and end. Portions of the traditional Christian liturgy are in ABA form, notably the *kyrie*, and also the *hosanna* portion of the *sanctus*. Staal (1984b, 1989) points out that many traditional musical forms have an ABA structure: rondo, minuet, and the sonata form derived from them. More complex palindromic forms are also found in the classical music repertoire, e.g. ABACABA was a common form of the rondo during the classical period, and Staal cites sequences as baroque as ABACABACABA and ABCDEDCBA in Bach's compositions. Interestingly, Senghas, Kita & Özyürek (2004) also reported the use of ABA structure to express relations of simultaneity in one third of their Nicaraguan Sign Language participants, regardless of cohort. In any event, the degree of center-embedding in ritualized human activity, as Staal observed with regard to more intricate, longer Vedic ritual performances, likely antedates and informs its use in language.

4.2. Animal ritual

There is also some limited evidence for simple ABA center-embedded structures in animal behavioral rituals, i.e. a central action bracketed by two identical or matching actions. For example, greeting rituals between male baboons (Smuts & Watanabe, 1990; Whitham & Maestipieri, 2003; Dal Pesco & Fischer, 2018; see also Colmenares, 1990, 1991) have a quasi-palindromic structure: one baboon approaches another with obvious intent, performs the greeting, and then retreats. The courtship dances of great crested grebes, as discussed by Staal (1984a), also show elements of this structure: the birds shake their heads, pick up water-weed, present it to each other, drop it, and then shake their heads again (Huxley, 1914:26ff). Interestingly, this behavioral bracketing of animal ritual sequences is mirrored by increased neuronal firing at the start and end of learned action sequences in the striatum of rats (Jog et al., 1999; Barnes et al., 2005) and the prefrontal cortex of macaques (Fujii & Graybiel, 2003).

4.3. Animal artificial grammar learning

Independent of Staal's claims, attempts have been made to ascertain whether other animals can learn the recursive grammatical rules necessary to interpret or produce center-embedded sequences. The finding of Gentner et al. (2006) that European starlings can learn center-embedded auditory strings of the form A^nB^n up to $n = 3$ (i.e. AAABBB) could also be accounted for by a finite-state grammar augmented with a counter (Rogers & Pullum, 2011:339). But center-embedding rules are unambiguously required to form or interpret palindromic sequences (e.g. ABCBA), and there is evidence that both songbirds (Abe & Watanabe, 2011) and macaques (Jiang et al., 2018) are able to recognize and complete such sequences.

5. Iterated learning and the emergence of center-embedding

In a series of non-linguistic iterated learning experiments, we set out to determine under what conditions center-embedded structures might emerge in sequences of pictorial icons. Two factors were manipulated. First, initial stimuli consisted of strings in which all icons were either entirely distinct, or else contained two non-adjacent identical icons, which could be understood as forming a dependency. This manipulation was designed to test the hypothesis that the existence of long-distance dependencies in the input stimuli might serve as a necessary precursor for the emergence of center-embedding (Pullum & Rogers, 2006; this is also consistent with the results of Abe & Watanabe, 2011, cited above in 4.3). Second, icon sequences were either presented simultaneously in a block, or sequentially one by one in RSVP mode. This was designed to test whether working memory constraints might facilitate or impede the emergence of center-embedding.

Thus far we have been able to establish two basic facts: (1) the repetition at a distance of icons in the initial input string does indeed facilitate the emergence of center-embedding, which (2) occurs to a greater degree under the additional working memory burden of sequential presentation. However, while center-embedded structures did appear over ten generations of transmission, they did not appear at levels greater than chance. Additionally, a follow-up experiment has shown that if input strings contain center-embedded structures (i.e. palindromes), this structure disappears in transmission, indicating that center-embedding is by itself neither a preferred nor an advantageous feature for the learning and transmission of strings, consistent with decades of psycholinguistic research.

6. Conclusion

We have resurrected and argued for the potential advantages of linguistic proposals for the evolution of language based on the scientific study of animal and human ritual. These proposals existed on the fringe of research in the 1980s, and have been largely forgotten today. To the extent that ritual has been a major focus of language evolution theorists, it has been primarily the domain of cultural and social anthropologists, or biologists and ethologists. Yet their conclusions are largely consonant with those of linguists: structural form predominates over content in human ritual (Lewis, 2018; Merker, 2005, 2009). Our goal here has been to help bring linguistic studies of ritual back into the mainstream of thought on language evolution, because it seems to us that they offer the possibility for the simultaneous evolution of crucial referential (i.e. arbitrary symbolic reference) and structural properties (i.e. recursion in the form of center-embedding) of language in tandem – but also possibly independently of each other – within the same general model. We therefore advocate a general research framework for addressing these questions going forward. We are focused on the question of recursion in the form of center-embedding and its emergence in human language, given its status as the apparent *sine qua non* of language evolution (Hauser, Chomsky & Fitch, 2002). However, an equivalent focus should be directed to the emergence of arbitrary reference and duality of patterning. Our suspicion and hope is that a general research agenda of this type could help to coordinate and align more closely the important research on language evolution of anthropologists, ethologists, biologists and linguists, to the overall benefit of the field as a whole.

References

Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, *14*(*8*), 1067.

Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., & Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*, *437*, 1158-1161.

Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes, (Ed.) *Cognition and the development of language* (pp. 1-61). New York: John Wiley and Sons.

Brandon, R. N., & Hornstein, N. (1986). From icons to symbols: Some speculations on the origins of language. *Biology and Philosophy*, *1*(*2*), 169-189.

Colmenares, F. (1990). Greeting behaviour in male baboons, I: Communication, reciprocity and symmetry. *Behaviour, 113*(*1-2*), 81-116.

Colmenares, F. (1991). Greeting, aggression, and coalitions between male baboons: Demographic correlates. *Primates, 32*(*4*), 453-463.

Dal Pesco, F., & Fischer, J. (2018). Greetings in male Guinea baboons and the function of rituals in complex social groups. *Journal of Human Evolution*, *125*, 87-98.

Fujii, N., & Graybiel, A. M. (2003). Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science*, *301*, 1246-1249.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*, 1204-1207.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? S*cience*, *298*, 1569-1579.

Huxley, J. (1914). *The courtship habits of the great crested grebe*. Jonathan Cape.

Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., & Wang, L. (2018). Production of supra-regular spatial sequences by macaque monkeys. *Current Biology, 28*(*12*), 1851-1859.

Jog, M. S., Kubota, Y., Connolly, C. J., Hillegaart, V., & Graybiel, A. M. (1999). Building neural representations of habits. *Science*, *286*, 1745-1749.

Lewis, J. (2018). Music before language: Observations from a hunter-gatherer's point of view. Plenary address, EvoLang XII, Toruń, Poland, April 19. *https://www.youtube.com/watch?v=xkwY84mfhNg&t=1889s*

Lorenz, K. (1977). *Behind the mirror: A search for a natural history of human knowledge.* New York, NY: Harcourt Brace Jovanovich.

Merker, B. (2005). The conformal motive in birdsong, music, and language: An introduction. In G. Avamzani, L. Lopez, S. Koelsch, & M. Majno (Eds.). *The Neuroscience and Music II: From Perception to Performance. Annals of the New York Academy of Sciences*, *1060*: 17-28.

Merker, B. (2009). Ritual foundations of human uniqueness. In S. Malloch, & C. Trevathen (Eds.). *Communicative Musicality: Exploring the Basis of Human Companionship* (pages 45-59). Oxford: Oxford University Press.

Poletiek, F. H., Fitz, H., & Bocanegra, B. R. (2016). What baboons can (not) tell us about natural language grammars. *Cognition*, *151*, 108-112.

Pullum, G. K., & Rogers, J. (2006). Animal pattern-learning experiments: Some mathematical background. Unpublished manuscript. Radcliffe Institute for Advanced Study/Harvard University.

Pulvermüller, F. (2014). The syntax of action. *Trends in Cognitive Sciences*, *18*(*5*), 219-220.

Rogers, J., & Pullum, G. K. (2011). Aural pattern recognition experiments and the subregular hierarchy. *Journal of Logic, Language and Information, 20*(*3*), 329-342.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science*, *305*, 1779-1782.

Smuts, B.B., & Watanabe, J.M. (1990). Social relationships and ritualized greetings in adult male baboons (Papio cynocephalus anubis). *International Journal of Primatology, 11*(*2*), 147-172.

Staal, F. (1979). The meaninglessness of ritual. *Numen*, 26(*1*), 2-22.

Staal, F. (1980). Ritual syntax. In M. Nagatomi, B.K. Matilal, J.M. Masson, & E. C. Dimock (Eds.). *Sanskrit and Indian Studies* (pp. 119-142). Dordrecht: Springer Netherlands.

Staal, F. (1984a). Ritual, mantras and the origin of language. In S. D. Joshi (Ed.), *Amṛtadhārā: Professor R.N. Dandekar Felicitation Volume* (pp. 403-425). Delhi: Ajanta Publications.

Staal, F. (1984b). The search for meaning: music, mathematics, and ritual. *The American Journal of Semiotics, 2*(*4*), 1-57.

Staal, F. (1989). *Ritual and mantras: Rules without meaning.* New York, NY: Peter Lang Publishing.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.

Tonna, M., Marchesi, C., & Parmigiani, S. (2019). The biological origins of rituals: An interdisciplinary perspective. *Neuroscience and Behavioral Reviews, 98*, 95-106.

Whitham, J. C., & Maestripieri, D. (2003). Primate rituals: The function of greetings between male guinea baboons. *Ethology*, *109,* 847-859.

IMPACT OF AUDITORY FEEDBACK ON BAT VOCAL DEVELOPMENT

ELLA Z. LATTENKAMP[1,2,*], MEIKE LINNENSCHMIDT[1], EVA MARDUS[1], SONJA C. VERNES[2], LUTZ WIEGREBE[1], MICHAEL SCHUTTE[1]

[*]Corresponding author: ella.lattenkamp@evobio.eu

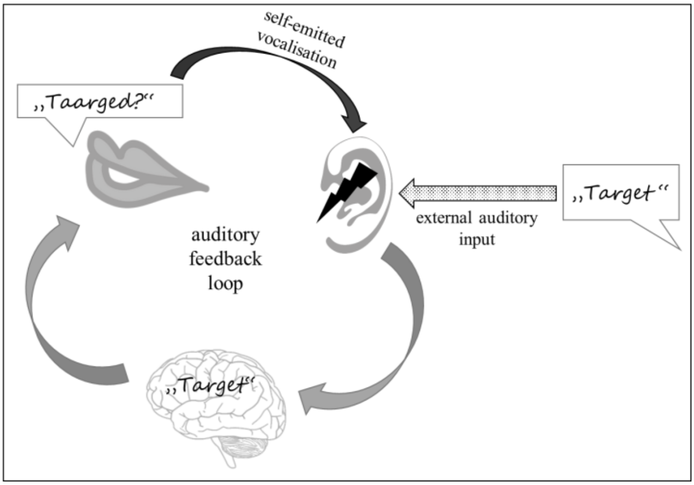
[1]AG Wiegrebe, Dept. Biology II, Ludwig-Maximilians University, Munich, Germany.

[2]Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

Speech acquisition in humans is based (1) on the perception of an external auditory target (e.g. adult vocalisations), and (2) the gradual modification of self-emitted vocalisations to match this acoustic target (figure 1) (Kuhl & Meltzoff, 1996). When hearing is severely impaired and auditory feedback is lacking (e.g. in deaf infants), normal adult vocalisations cannot be acquired (Oller & Eilers, 1988). Instead, the vocal development of deaf infants shows variations such as reduced inventories for consonants, vowels and syllables, reduced articulation space, and atypical temporal and coordinative sound production (Hudgins & Numbers, 1942; Brannon, 1966; Clement & Koopmans-van Beinum, 1995; VanDam *et al.*, 2015). Although speech and spoken language acquisition are uniquely human traits, the disrupting effect of lacking auditory feedback is not as rare in the animal kingdom. All species that are capable of vocal learning, i.e. able to acquire novel vocalisations through auditory input, should be impacted by a disruption of sensory input or auditory feedback (figure 1). Significant variations of normal adult vocalisations are expected to occur in any vocal learning species that is deprived of auditory feedback during development (as has been shown in songbirds (e.g. Konishi, 1965; Konishi, 2004)).

Bats have been highlighted as promising mammalian model organisms for the study of vocal learning (Knörnschild, 2014; Vernes & Wilkinson, 2019). Especially for the pale spear-nosed bat, *Phyllostomus discolor,* several indications of vocal learning have previously been published (Esser & Schmidt, 1989; Esser, 1994). Moreover, adult *P. discolor* have been shown to possess a rich vocal repertoire (Lattenkamp *et al.*, 2019) and to be flexible in the spectro-temporal domain of their vocalisations (Lattenkamp *et al.*, 2018). In order to provide further evidence for the vocal learning capacity of *P. discolor*, we investigated the impact of auditory input and feedback on their vocal development. Complementing previous isolation studies in this species (i.e. exclusive disruption of external auditory input), which focused on the effects on a single vocalisation type (Esser, 1994), we have now investigated the effect of deafening (i.e. disruption of auditory input and feedback) on their full vocal repertoire. Comparisons between the vocalisations of severely hearing-impaired and normal-hearing pups and adult bats allow us to assess the impact of auditory feedback on the vocal development of this species. The hearing impairment led to a significant increase in vocal activity and call amplitude as well as an increase in several other spectro-temporal call parameters. These results present an important contribution to the discussion of the status of *P. discolor* as a vocal learner and the role of this species for the study of vocal learning. This work serves as a basis for further research using the pale spear-nosed bat as a mammalian animal model for vocal learning and contributes to comparative studies on hearing impairment across species.

Figure 1. Schematic of an auditory feedback loop underlying speech acquisition and vocal learning.



In a functional auditory feedback loop the auditory target, e.g. an adult vocalisation (dotted arrow), is perceived and memorised. The self-emitted vocalisation of the target sound is then perceived again (black arrow) and compared to the memorised target. While isolation studies solely prevent the perception of external auditory targets, the disruption of the auditory feedback loop by deafening (black lightning symbol) also disrupts the perception of self-emitted vocalisations and thus the possibility for adjustment towards an internal target.

Acknowledgements

This work was funded by a Human Frontiers Science Program (HFSP) Research grant (RGP0058/2016) awarded to L.W. and S.C.V. The authors want to thank Stephanie Shields and the LMU workshop for help with constructing parts of the setup used in this study and Sabrina Schenk for her excellent bat care.

References

Brannon, J. (1966). The speech production and spoken language of the deaf. *Language and Speech*, 9, 127-36.

Clement, C. J., & Koopmans-van Beinum, F. J. (1995). Influence of lack of auditory feedback: vocalizations of deaf and hearing infants compared. *IFA Proceedings 19*, 25-37.

Esser, K.-H. (1994). Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat *Phyllostomus discolor*. *Neuroreport*, 5: 1718-20.

Esser, K.-H., & Schmidt, U. (1989). Mother-infant communication in the lesser spear-nosed bat *Phyllostomus discolor* (Chiroptera, phyllostomidae): Evidence for acoustic learning. *Ethology*, 82(2), 156-168.

Hudgins, C. V., & Numbers, F. C. (1942). An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*, 25, 289-392.

Knörnschild, M. (2014). Vocal production learning in bats. *Current Opinion in Neurobiology*, 28, 80-85.

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Zeitschrift für Tierpsychologie*, 22, 770-83.

Konishi, M. (2004). The role of auditory feedback in birdsong. *Annals of the New York Academy of Sciences*, 1016, 463-75.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. *Journal of the Acoustical Society of America*, 100, 2425–2438.

Lattenkamp, E. Z., Vernes, S. C., Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *Journal of Experimental Biology*, 211, jeb180729.

Lattenkamp, E. Z., Shields, S. M., Schutte, M., Richter, J., Linnenschmidt, M., Vernes, S. C., Wiegrebe, L. (2019). The vocal repertoire of pale spear-nosed bats in a social roosting context. *Frontiers in Ecology and Evolution*, 7, 16.

Oller, K. D., & Eilers, R. E. (1988). The role of audition in infant babbling. *Child Development*, 59(2), 441-449.

VanDam, M., Oller, K., Ambrose, S. E., Gray, S., Richards, J. A., Xu, D., Gilkerson, J., Silbert, N. H., & Moeller, M. P. (2015). Automated vocal analysis of children with hearing loss and their typical and atypical peers. *Ear Hear*, 36, e146-52.

Vernes, S. C., & Wilkinson, G. S. (2019). Behaviour, biology, and evolution of vocal learning in bats. *Philosophical Transactions of the Royal Society of London, Series B*, 1–31.

USING SPATIAL VISUALIZATIONS AND REAL-WORLD SOCIAL NETWORKS TO UNDERSTAND LANGUAGE EVOLUTION AND CHANGE

LI LEI*[1,2], LIMOR RAVIV[2], and PHILLIP ALDAY[2]

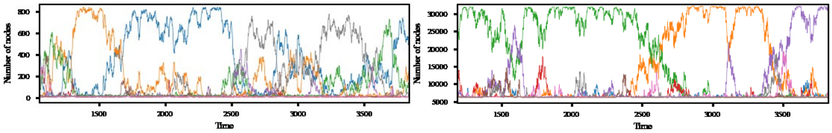
*Corresponding Author: leili@science.ru.nl
[1]Radboud University, Nijmegen, The Netherlands
[2]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Over the last decade, computational modelling has emerged as a powerful tool for studying language evolution and change in social networks. The main aim of these studies is to explore the propagation of variants in different networks. In particular, Fagyal, Swarup, Escobar, Gasser, & Lakkaraju (2010) proposed a degree-biased voter model (DBVM) to investigate the role of hubs and loners in the community: different linguistic variants could spread in a bi-directional closed network, with higher in-degree agents (i.e., with more possible connections) having a higher probability to be chosen as interaction partners. The model revealed that while hubs promoted the spread of variants in the network and facilitated convergence, the loners (i.e., the more isolated agents) played a key role in language change: loners seem to serve as “variant-keepers” and maintain their unique variants despite the prevalence of another norm. By doing so, they seem to prevent the loss of variants and eventually reintroduce these variants to the entire community.

However, the network structure used in Fagyal et al. (2010) was based on a simulated artificial network, which does not necessarily adhere to all properties of real-world networks. Moreover, the exact role of loners was evaluated based on extreme and indirect manipulations, such as removing loners from the network altogether. In this study, we extend the model of Fagyal et al. (2010) by (a) incorporating more realistic social networks, and (b) visualizing the process of language diffusion to examine the exact role of loners.

We first replicated Fagyal et al. (2010)’s model using the same parameters, and obtained similar diffusion results (Figure 1a): over time, competing variants (represented by different colours) alternately became the norm. We then visualized the process of language diffusion spatially across multiple simulations (Figure 2). This method directly demonstrated the unique role of the loners: At first (*t*=0; Figure 2a), the eight possible variants were uniform and randomly distributed. After multiple interactions (*t*=150; Figure 2b), one variant (in green) spread throughout the network and became the norm. Nevertheless, some agents maintained less dominant variants (in blue), and the loners kept even rarer variants (in purple). Later (*t*=300; Figure 2c), those variants became the norm, while rarer variants were still maintained.

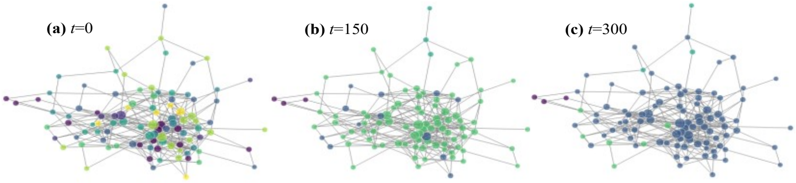
We then tested the model on a large-scale, real-world "who-trust-whom" online social network, which was extracted from the review website Epinions (Richardson, Agrawal & Domingos, 2003). We found the same diffusion dynamics in the real-world network and in the artificial network (Figure 1b). Finally, we scaled down the Epinions network and manipulated the numbers of edges (i.e., the number of possible connections). The results showed similar trends, i.e., alternating linguistic norms over time. Importantly, we found that increasing the number of connections led to a longer and smoother fixation of linguistic variants. This result has important implications for the process of cultural transmission and language evolution, and suggests that language change is slower in a dense network. It is also in line with claims of greater variability and less conventionalization in emerging sign languages from dense communities (Meir, Israel, Sandler, Padden & Aronoff, 2012), and is in line with the empirical work showing that norms are easier to maintain in a high-density community (Milroy, 1987). In on-going work, we are using more common communication networks (e.g., Twitter) to further reduce the gap between simulated networks, online networks and real-world social networks.



(A) Original atificial network

(B) Real-world Epinions network

Figure 1. Changes in the prominence of linguistic variants over time in the original network (A) and in the real-world Epinions network (B)



(a) *t*=0

(b) *t*=150

(c) *t*=300

Figure 2. (a) *t*=0, eight linguistic variants are assigned uniformly and randomly to agents; (b) *t*=150, green variants are dominant in the network; (c) *t*=300, blue variants are dominant in the network.

Acknowledgments

We thank three anonymous reviewers and the audiences at the Symposium on Representations, Usage and Social Embedding in Language Change in Manchester in 2019 for their comments.

References

Fagyal, Z., Swarup, S., Escobar, A. M., Gasser, L., & Lakkaraju, K. (2010). Centers and peripheries: Network roles in language change. Lingua, 120(8), 2061-2079.

Meir, I., Israel, A., Sandler, W., Padden, C. A., & Aronoff, M. (2012). The influence of community on language structure: Evidence from two young sign languages. Linguistic Variation, 12(2), 247-291.

Milroy, L. (1987). Language and social networks. Oxford: Basil Blackwell.

Richardson M., Agrawal R., & Domingos P. (2003). Trust management for the semantic web. In D. Fensel, K. Sycara & J. Mylopoulos (Eds.), Proceedings of the Second International Semantic Web Conference (ISWC 2003) (pp. 351-368). Berlin, Heidelberg: Springer.

CALL COMBINATIONS IN CHIMPANZEES

MAËL LEROUX*[1,2,3], KATIE E. SLOCOMBE[4], ANNE M. SCHEL[5], KLAUS ZUBERBÜHLER[2,3,6,7] AND SIMON W. TOWNSEND[1,3,8]

*Corresponding Author: mael.leroux@uzh.ch
[1] Department of Comparative Linguistics, University of Zürich, Zürich, Switzerland
[2] Budongo Conservation Field Station, Masindi, Uganda
[3] Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zürich, Zürich, Switzerland
[4] Department of Psychology, University of York, York, UK
[5] Animal Ecology, Utrecht University, Utrecht, Netherlands
[6] Department of Comparative Cognition, University of Neuchâtel, Neuchâtel, Switzerland
[7] School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK
[8] Department of Psychology, University of Warwick, Warwick, UK

A key feature of language often argued to distinguish it from other animal communication systems is its compositionality: meaningful words are combined together into larger structures with derived meaning (Hurford, 2011). Despite its significance, we know very little regarding how unique this capacity is to humans or the evolutionary progression of this trait. By searching for core features of language, such as its compositionality, in the communication systems of our closest-living relatives, particularly the apes, it is possible to shed light on whether the components of language are *de novo* evolved traits in humans or whether they have their origins rooted in the primate lineage. Using a high-density data approach, we are currently investigating the forms of vocal combinatorial structures in chimpanzees (*Pan troglodytes schweinfurthii*). Preliminary data collected in the Budongo forest, Uganda, indicates that call combinations represent 13% of the total vocal production in chimpanzees and the context of production is a good predicator of the call types combined. Social contexts, in particular, seem to elicit the highest proportion of call combinations, supporting the hypothesis that sociality might be an important driver for the evolution of combinatorial communication systems (Collier, Townsend, & Manser, 2017). Moreover, predator presentations revealed that chimpanzees combine calls (alarm-hoos with waa-barks) when encountering a snake, specifically when other individuals are present, potentially to recruit group members in a dangerous situation. Ongoing playback experiments aim to further investigate the meaning and function of this call combination and to test whether chimpanzees process it

as a compositional-like structure. Preliminary analyses of responses to playbacks suggests the alarm-hoo-waa-bark combination is meaningful to receivers and, critically, related to the individual meaning of the comprising calls.
Ultimately, this work will help elucidate whether one of the core building blocks of language, compositionality, also exists in the communication systems of our closest relative and therefore is evolutionarily more ancient.

Acknowledgements

We are grateful to the staff of the Budongo Conservation Field Station for supporting us with this study, and to the Royal Zoological Society of Scotland for providing core funding. We also thank the Ugandan Wildlife Authority, the Ugandan National Council for Science and Technology and the President's Office for permission to conduct this research. We thank our field assistants, Bosco Chandia and Denis Lomoro for helping with data collection.

References

Collier, K., Townsend, S. W., & Manser, M. B. (2017). Call concatenation in wild meerkats. *Animal Behaviour*, *134*, 257–269.
Hurford, J. (2011). *The origins of grammar: Language in the light of evolution II*. Oxford: Oxford Univerity Press.

COMMUNITIES OF DIFFERENT SIZE CREATE DIFFERENT CATEGORIZATION SYSTEMS

SHIRI LEV-ARI[*1]

[*]Corresponding Author: shiri.lev-ari@rhul.ac.uk
[1]Dept. of Psychology, Royal Holloway University of London, UK

Categorization is the foundation of many cognitive functions, and the way in which we categorize the world is informed by the language we speak. Languages, however, differ in the granularity of the categories they encode, and the source of these cross-linguistic differences is poorly understood. Prior research has shown that larger communities have larger phonological inventories in both human (Hay & Bauer, 2007) and non-human animals (e.g., Freeberg, 2006; McComb & Semple, 2005) as well as create more systematic languages (Lupyan & Dale, 2010; Raviv, Meyer & Lev-Ari, 2019). This paper tests whether community size and density can also influence the granularity and structure of semantic categories, a domain that at the interface between language and cognition. As such, this study will not only add to our understanding of why and how languages evolved to have the forms they have, but it will also have implication for linguistic and non-linguistic cognitive performance, as the way a language encodes a category has implications for memory, attention, and even low level perception (e.g., Roberson, Davidoff, Davies & Shapiro, 2005; Winawer et al., 2007).

The positive association between community size and the size of the phonological inventory suggests that larger communities are likely to develop more granular categories, that is, categories with more sub-divisions than those of smaller communities. Additionally, larger communities' tendency to develop more systematic symbols suggests that they might also develop categorization systems with more informative structure that leads to better performance. In addition, we tested whether community density plays a similar role since sparsity, similarly to size, can influence input variability and diffusion mechanisms.

The proposal that community structure can influence the community's categorization system was tested with simulations in which populations of either 100 or 200 agents communicated about a 20x20 meaning space for 50,000 rounds.

Populations were generated using the barabasi_albert_graph in the Python package, Networkx (Barabási & Albert, 1999). Density was manipulated using the m parameter (m=20 vs 50). In each round, each agent communicated to someone in their network about a randomly selected meaning. Success was measured by the distance of the comprehended meaning from the intended one. Agents selected a label to produce by searching their history for the closest and most successful label, and when none was available, created a new label by randomly combining 3 phonemes. Partners interpreted the label according to their past experience with the label (weighting tokens by their success). If a label or a meaning was not used for at least 500 interactions, the agent forgot it.

Results showed that larger communities divided the meaning space into more categories (Fig 1a). Furthermore, these categories were more balanced in size (Fig 1b). As a consequence, larger communities communicated more successfully (Fig 1c). The effects of density were smaller in magnitude and less consistent.

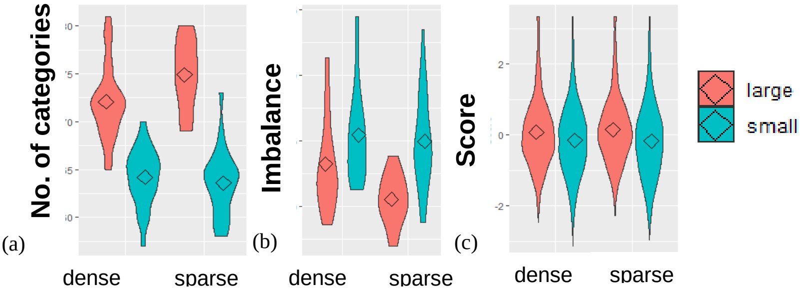


Figure 1. Properties of the categorization system as a function of community size and community density.

Further analyses revealed that it was larger communities' greater diffusion constraints that led to their superior behavior. The statistical analyses showed that: Labels were less likely to spread in larger communities, and even when they did, there was lower agreement on their meaning. This lesser agreement led categories to narrow their meanings, which, in turn, enabled their maintenance and the creation of more granular and balanced categories.

Lastly, matching data from the World Color Survey with population size from Ethnologue (controlling for language family) showed that larger populations have more color terms, providing initial real-life support for our findings.

This study shows that community structure influences the categorization system, such that larger communities create more granular and better structured categories that can support better communication. Moreover, the study shows how greater constraints – diffusion barrier – can ultimately improve a system. As greater granularity can add processing costs, future research should examine how larger communities might balance out the costs (e.g., more distinct labels).

References

Barabási, A. L., & Albert, R. (1999) Emergence of scaling in random networks. *Science 286,* 5439, 509-512.

Lupyan, G., and Dale, R. (2010) Language Structure Is Partly Determined by Social Structure. *Plos One* 5, e8559.

Freeberg, T. M. (2006) Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychological Science, 17, 7,* 557-561.

Hay, J. & Bauer, L. (2007) Phoneme inventory size and population size. Language 83, 388–400.

McComb, K. & Semple, S. (2005) Coevolution of vocal communication and sociality in primates. *Biology Letters 1,* 381 –385.

Raviv, L., Meyer, A.S. & Lev-Ari, S. (2019) Larger communities create more systematic languages. *Proceedings of The Royal Society B 286*, 201912.

Roberson, D., Davidoff, J., Davies, I. R., & Shapiro, L. R. (2005) Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology 50, 4*, 378-411.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007) Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences 104, 19*, 7780-7785.

A POINT IS NOT A POINT IS NOT A POINT: RE-INTERPRETING THREE BASIC KINDS OF POINTING COMPREHENSION

HEIDI LYN[*1] and JENNIE CHRISTOPHER[1,2]

[*]Corresponding Author: hlyn@southalabama.edu
[1]Department of Psychology, University of South Alabama, Mobile, AL USA
[2] Department of Psychology, University of Southern Mississippi, Hattiesburg, MS USA

The comprehension of human points has been a topic of debate over many years for those studying language evolution. Pointing is seen as a precursor behavior to referential communication and as such as a potential focus for comparative studies. Unfortunately, because humans seem to comprehend all pointing types as similarly referential, many researchers assume that animals similarly treat all points as the same. Therefore, many studies treat all point types as equally interesting and important for studying language evolution, although there is considerable evidence that suggests otherwise (see Lyn, 2010; Mulcahy & Hedge, 2012; Osborne & Mulcahy, 2019 for examples). Such methodological imprecision confuses categorically different concepts and can muddy the interpretive waters and inhibit scientific progress (e.g. Clark, Elsherif, & Leavens, 2019). We suggest here and will provide preliminary evidence that only one of the many types of human points is valuable for comparative study of communicative processes, and human language evolution in particular. Further, we suggest that only one hypothesis regarding nonhuman comprehension of this point – what we are calling the Social Interaction Point Hypothesis – explains the data so far collected.

To illustrate our framework, we will present and compare new and published canine and ape performance on three main point types, specifically as tested within the object-choice task, in which the participant is presented with an array of options, the correct choice being indicated by a point:

Proximal-Proximal (PP): The choice items are close together and the point is close to the correct item – usually within a few inches.

Proximal-Distal (PD): The choice items are close together, but the point is further away – usually more than 10-12 inches (also called Abject Object Choice (Mulcahy & Hedge, 2012)

Distal-Distal (DD): The choice items are further apart (usually more than 3 feet), and the point is therefore necessarily further away.

There are further complications to pointing behavior, such as ipsilateral vs contralateral pointing, asymmetric pointing, etc. But our framework neatly encompasses the current findings from those point types as well.

Our data, as well as the preponderance of evidence from the past 20 years of pointing studies, suggest that both PP and DD point following can be explained by simple social learning or by associative learning mechanisms and therefore do not require the subjects to recognize the communicative nature of the gesture. Specifically, researchers over the past 20 years have shown that when the pointing gesture is proximal to the item, the only mechanism required to successfully complete the task is stimulus enhancement (Osborne & Mulcahy, 2019; Povinelli, Bierschwale, & Ĉech, 1999). Similarly, our data supports earlier findings show that DD point following can be achieved by simpler mechanisms, such as associative learning or simple heuristics (Call, 2001; Miklosi & Soproni, 2006; Povinelli, Bierschwale, & Čech, 1999).

In contrast, the PD points are likely the only point types that require triadic perspective taking and an understanding of the communicative nature of the point, which explains the decreased success in our apes and dogs as well as supporting the findings of other researchers in apes (e.g. Clark et al., 2019; Leavens, Hopkins, & Bard, 2005). Therefore, to any researcher interested in exploring the potential for point following as a precursor to language, the PD point types are the only types of interest.

This framework explains why many attempts to place point following within the framework of human language evolution have failed. The domestication hypothesis (Hare et al., 2010), for instance, failed to account for systematic methodological differences across species. The mechanistic perspective (see Povinelli, Bierschwale, & Čech, 1999) failed to note that PD point forms (and a few others) do not allow for non-communicative learning mechanisms. Instead, the data on communicative points all support one conclusion: the best predictor of success is human social contact. We and others have reported, but not formalized this more accurate representation of the data, what we suggest calling the Social Interaction Point Hypothesis (e.g. Lyn, Russell, & Hopkins, 2010; Udell, Spencer, Dorey, & Wynne, 2012). This hypothesis, is also scientifically preferable because instead of separating communicative points as an anomalous gesture, it returns them to the context of the evolution of gestural communication more broadly defined.

References

Call, J. (2001). Chimpanzee social cognition. *Trends in Cognitive Sciences*, *5*(369–405).

Clark, H., Elsherif, M. M., & Leavens, D. A. (2019). Ontogeny vs. Phylogeny in Primate/Canid Comparisons: A Meta-analysis of the Object Choice Task. *Neuroscience & Biobehavioral Reviews*, *105*(January), 178–189. https://doi.org/10.1016/j.neubiorev.2019.06.001

Hare, B., Rosati, A., Kaminski, J., Bräuer, J., Call, J., & Tomasello, M. (2010). The domestication hypothesis for dogs' skills with human communication: A response to Udell et al. (2008) and Wynne et al. (2008). *Animal Behaviour*, *79*(2), e1–e6. Retrieved from 10.1016/j.anbehav.2009.06.031

Lyn, H. (2010). Environment, methodology, and the object choice task in apes: evidence for declarative comprehension and implications for the evolution of language. *Journal of Evolutionary Psychology*, *8*(4), 333–349. https://doi.org/10.1556/JEP.8.2010.4.3

Lyn, H., Russell, J. L., & Hopkins, W. D. (2010). The impact of environment on the comprehension of declarative communication in apes. *Psychological Science*, *21*(3), 360–365. https://doi.org/10.1177/0956797610362218

Miklosi, A., & Soproni, K. (2006). A comparative analysis of animals' understanding of the human pointing gesture. *Animal Cognition*, *9*, 81–93. https://doi.org/10.1007/s10071-005-0008-1

Mulcahy, N. J., & Hedge, V. (2012). Are great apes tested with an abject object-choice task? *Animal Behaviour*, *83*(2), 313–321. https://doi.org/10.1016/j.anbehav.2011.11.019

Osborne, T., & Mulcahy, N. J. (2019). Reassessing shelter dogs' use of human communicative cues in the standard object-choice task. *Plos One*, *14*(3), e0213166. https://doi.org/10.1371/journal.pone.0213166

Povinelli, D. J., Bierschwale, D. T., & Ĉech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, *17*(Pt 1), 37–60. https://doi.org/10.1348/026151099165140

Povinelli, D. J., Bierschwale, D. T., & Čech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, *17*(1), 37–60. https://doi.org/10.1348/026151099165140

Udell, M. A. R., Spencer, J. M., Dorey, N. R., & Wynne, C. D. L. (2012). *Human-Socialized Wolves Follow Diverse Human Gestures … And They May Not Be Alone* . 97–117.

DOES A UNIVERSAL HIERARCHICAL STRUCTURE UNDERLIE WORD ORDER TYPOLOGY?

ALEXANDER MARTIN*[1], PATRICK KANAMPIU[2], KLAUS ABELS[3], DAVID ADGER[4], and JENNIFER CULBERTSON[1]

*Corresponding Author: alxndr.martin@gmail.com
[1]Centre for Language Evolution, University of Edinburgh, United Kingdom
[2]Tharaka University College, Kenya
[3]University College London, United Kingdom
[4]Queen Mary University of London, United Kingdom

Theoretical syntacticians have argued that a universal underlying hierarchical structure has shaped the evolution of noun phrase word order: adjectives are structurally closest to the noun, then numerals, then demonstratives (Cinque, 2005; Abels & Neeleman, 2012). This structure can explain why, typologically, orders that transparently reflect this hierarchical structure (termed *homomorphic* orders, e.g., Dem-Num-Adj-N, Num-N-Adj-Dem, etc.) overwhelmingly outnumber orders that do not (e.g., N-Dem-Num-Adj) (Dryer, 2018). A recent series of experiments appears to support this hypothesised structure, showing that learners' inferences about word order in a new language reflect this structure, rather than linear order (Culbertson & Adger, 2014; Martin, Ratitamkul, Abels, Adger, & Culbertson, in press; Martin, Abels, Adger, & Culbertson, 2019). Participants were taught a version of their native language (or an artificial language) in which the order of a single modifier (Adj, Num or dem) relative to the noun was swapped (e.g., English speakers learned that "green car" was produced "car green" in the new language). Participants then had to guess the relative order of multiple modifiers (e.g., "two green cars"). Participants consistently inferred orders which reflected not their native language order transposed (i.e., "cars two green"), but instead the underlying hierarchical structure in this domain (i.e., "cars green two").

However, these previous experiments targeted English and Thai speakers, whose native languages are homomorphic (Dem-Num-Adj-N and N-Adj-Num-Dem respectively). From these populations, it is thus impossible to determine whether there is a universal preference for homomorphism—causally connected to typology—or whether instead the results reflect structural transfer from participants' native language. If there is a universal homomorphism bias, then even speakers of a language whose word order is *not* homomorphic should infer homomorphic word order in an artificial language. The present project aims to test this prediction, by comparing word order preferences from English-speaking participants with those from a population whose native language, Kîîtharaka, is not homomorphic—specifically, N-Dem-Num-Adj (Kanampiu, 2017).

We used an artificial language learning task to test our hypothesis. Participants were instructed that they would be learning part of a new language called Nápíjò. In actuality, Nápíjò was an artificial language composed of three nouns and a series of modifiers, following the materials design described by Martin et al. (2019). Contrary to previous studies, however, our task required participants to produce noun phrases in the artificial language. Participants were taught word meanings by hearing labels accompanying simple cartoon scenes. Objects were depicted on a table positioned in front of a girl. Participants first learned the object names by seeing greyscale images of the objects. Participants then saw modified objects (e.g., a red feather, two mugs),[1] and heard the noun *followed* by the modifier label (in the version for English speakers) or *preceded* by the modifier label (for Kîîtharaka speakers). That is, in Nápíjò, all modifiers were on the opposite side of the noun from participants' native language. Once participants had learned the individual word meanings, and how nouns combined with single modifiers in the language, they were presented with visual stimuli that involved a colour *and* either a numeral (e.g., a group of three red feathers) or a demonstrative (e.g., the cartoon girl pointing to a red feather in front of her). They were asked how to describe these scenes—which they had not seen before—by producing phrases themselves in Nápíjò. Responses were automatically recorded and coded offline.

We tested 60 English-speaking participants who reported no knowledge of a language with postnominal modifiers (30 in each condition). We found a strong preference for homomorphic orders for both Dem+Adj (mean: 82.6%, $p < 0.001$) and Num+Adj combinations (mean: 67.5%, $p < 0.01$) (though the preference was stronger for the former than for the latter ($p < 0.05$)).

We are currently running the Nápíjò experiment with monolingual Kîîtharaka-speaking participants in the Tharaka region of Eastern Kenya. Because Kenya has a strong multilingual policy in its schools (both English and Swahili are compulsory), the population that we are targeting is an older illiterate population in rural areas with little exposure to and no fluency in languages like English. Data collection is ongoing and will be completed over the next six weeks, and comparison with our English results will allow us to say if those speakers, who have a lifetime of experience with a non-homomorphic language, will prefer homomorphic orders in Nápíjò. If so, this will provide strong support for the role of a universal underlying hierarchical structure in shaping how noun phrase word order has evolved. If Kîîtharaka speakers rather prefer non-homomorphic orders (likely a structural transfer from experience with their own language), this result will challenge our understanding of universal language structures and the pressures that shape linguistic systems over time.

---

[1] Demonstrative meanings were taught by showing the girl pointing at a proximal or a distal object.

References

Abels, K. & Neeleman, A. (2012). Linear asymmetries and the LCA. *Syntax*, *15*(1), 25–74.

Cinque, G. (2005). Deriving Greenberg's Universal 20 and Its Exceptions. *Linguistic Inquiry*, *36*(3), 315–332. JSTOR: 4179327

Culbertson, J. & Adger, D. (2014). Language learners privilege structured meaning over surface frequency. *PNAS*, *111*(16), 5842–5847.

Dryer, M. S. (2018). On the Order of Demonstrative, Numeral, Adjective and Noun. *Language*.

Kanampiu, P. N. (2017). *The syntax of the determiner phrase in Kîîtharaka, a Bantu language spoken in Kenya* (Master's thesis, Chuka University).

Martin, A., Abels, K., Adger, D., & Culbertson, J. (2019). Do learners' word order preferences reflect hierarchical language structure? In A. K. Goel, C. M. Seifart, & C. Freksa (Eds.), *Proceedings of the 41st Annual Meeting of the Cognitive Science Society* (pp. 2303–2309). Montreal, Canada: Cognitive Science Society.

Martin, A., Ratitamkul, T., Abels, K., Adger, D., & Culbertson, J. (in press). Cross-lingusitic evidence for cognitive universals in the noun phrase. *Linguistics Vanguard*.

ADDRESSING THE “VOCAL” IN VOCAL LEARNING

PEDRO TIAGO MARTINS*[1,2] and CEDRIC BOECKX[1,2,3]

*Corresponding Author: pt.martins@ub.edu
[1]Section of General Linguistics, University of Barcelona, Barcelona, Spain
[2]University of Barcelona Institute of Complex Systems, Barcelona, Spain
[3]Catalan Institute for Research and Advanced Studies, Barcelona, Spain

1. Vocal learning is usually about the learning

Vocal learning is hard to define as a phenotype. It is not known how pervasive it is in the animal kingdom, but it is generally considered to be present in at least three families of birds (songbirds, parrots, and hummingbirds), and in some mammals, which include cetaceans (Janik, 2014), pinnipeds (Ravignani et al., 2016), bats (Vernes, 2017) and elephants (Stoeger & Manger, 2014), and humans. This is understood to refer to (at least) vocal production learning, one of the 3 subtypes in what has become an well established typology, after (Janik & Slater, 2000):

vocal comprehension learning ability to associate a signal with a behavioral response

vocal usage learning ability to learn the context in which a vocalization can be used

vocal production learning ability to significantly modify vocalizations on the basis of experience

The crucial differences between these types of vocal learning have to do with *what is learned.* The first two types (*comprehension* and *usage*) pertain to the association of existing signals with new contexts, while the third kind pertains to the learning of the signals themselves. This could involve either a modificiation, as long as it can be shown that such a modification is learned from experience (e.g. auditory), or the production of novel vocalizations altogether. It is sometimes considered that only the production of novel calls can be considered to be vocal learning (Fitch, 2010). The questions of what counts as novel, what is learned, and what are the minimal units of animal calls for each species are not trivial, and they very often influence which vocal learning category a behavior will be ascribed to. Naturally, these questions inform much of the work on vocal learning.

2. The vocal in vocal learning

However, there is an important aspect of vocal learning that is rarely discussed or defined with little justification, which has to do not with the learning of signals but with their source. In other words, what "vocal" actually is. In a sense, this should not be important, if what is interesting about the phenotype is the learning. But in practice this issue goes beyond terminology: it rigidly contrains what different researchers count as a behavior to be studied, and has ramifications that go from phylogeny to the neurobiology and beyond. We identify different possibilities, which go from most limitative to most permissive:

- phonatory muscles: ability to produce sound by controling the larynx (e.g. humans) or syrinx (e.g. songbirds) (Elemans et al., 2015)
- any part of vocal tract: ability to produce sound by controling structures that could include the phonatory muscles but could be limited to the structures in the upper vocal tract, such as orangutan whistling (Lameira et al., 2013)
- any orofacial structure: ability to produce sound by using orofacial structures that might be outside of the vocal tract proper, such as the nasal tract in elephants (Stoeger & Manger, 2014)
- "artifical extension" of the vocal tract: ability to modify sounds by using external means outside of orofacial structures, such as hands or foreign objects, (e.g. orangutan (Hardus, Lameira, Schaik, & Wich, 2009))
- shortand for sound volitionally produced by animals, regardless of anatomy but usually assumed to involve any or all the means listed above

We contend that there are no *apriori* reasons for choosing one option over the other, and that each of them yield completely different approaches to vocal learning that have ramifications beyond behavior, and make classification and comparative work more challenging. One clear case is neurobiology: it is generally assumed that vocal learning involves a direct forebrain projection to the phonatory muscles (Fitch, 2010; Jarvis, 2007), and this idea has been used as a guideline for all species (Petkov & Jarvis, 2012). Species not shown to volitionally produce sounds using these anatomical and brain structures are often considered to be non-vocal learners, *regardless of their learning abilities pertaining to conspeficic calls*. (Ghazanfar, Liao, & Takahashi, 2019).

We discuss the implications of this state of affairs for the classifcation of different species in vocal learning, for the study of its neurobiology, and for devising evolutionary theories of this phenotype and other, closely related ones, with language as an important case.

References

Elemans, C., Rasmussen, J., Herbst, C., Düring, D., Zollinger, S., Brumm, H., Srivastava, K., Svane, N., Ding, M., Larsen, O., Sober, S., & Švec, J. (2015). Universal mechanisms of sound production and control in birds and mammals. *Nature Communications*, *6*(1), 8978.

Fitch, W. T. (2010). *The evolution of language.* Cambridge: Cambridge University Press. (OCLC: 703204991)

Ghazanfar, A. A., Liao, D. A., & Takahashi, D. Y. (2019). Volition and learning in primate vocal behaviour. *Animal Behaviour*.

Hardus, M. E., Lameira, A. R., Schaik, C. P. V., & Wich, S. A. (2009). Tool use in wild orang-utans modifies sound production: A functionally deceptive innovation? *Proceedings of the Royal Society of London B: Biological Sciences*, *276*(1673), 3689-3694.

Janik, V. M. (2014). Cetacean vocal learning and communication. *Current Opinion in Neurobiology*, *28*, 60-65.

Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*(1), 1–11.

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. *Journal of Ornithology*, *148*(1), 35–44.

Lameira, A. R., Hardus, M. E., Kowalsky, B., de Vries, H., Spruijt, B. M., Sterck, E. H. M., Shumaker, R. W., & Wich, S. A. (2013). Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-ended call repertoire: A replication and extension. *The Journal of the Acoustical Society of America*, *134*(3), 2326-2335.

Petkov, C. I., & Jarvis, E. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*.

Ravignani, A., Fitch, W. T., Hanke, F. D., Heinrich, T., Hurgitsch, B., Kotz, S. A., Scharff, C., Stoeger, A. S., & de Boer, B. (2016). What Pinnipeds Have to Say about Human Speech, Music, and the Evolution of Rhythm. *Frontiers in Neuroscience*, *10*.

Stoeger, A. S., & Manger, P. (2014). Vocal learning in elephants: Neural bases and adaptive context. *Current Opinion in Neurobiology*, *28*, 101-107.

Vernes, S. C. (2017). What bats have to say about speech and language. *Psychonomic Bulletin & Review*, *24*(1), 111-117.

TOWARDS A COMPARATIVE APPROACH TO (SYNTACTIC) HIERARCHY

DAIKI MATSUMOTO*[1]

*Corresponding Author: matsumoto.daiki.24x@st.kyoto-u.ac.jp
[1]Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

It is a well known fact that the human language syntax has a hierarchical nature (Chomsky 1957; see also Berwick & Chomsky 2016). The process of combining words into phrases and sentences proceeds in a recursive, non-Markovian fashion, instead of forming a string. Given that this is one of the most striking properties of the human language syntax as many researchers claim (see Fitch & Hauser, 2004; Fitch, 2010 among many others), the mechanism underlying this property should be treated as an *explanandum*. Generative grammarians claim that "Merge," an elementary combinatoric syntactic operation, neatly captures this property. The definition of Merge is shown in (1) below (see Chomsky, 2010; Berwick & Chomsky, 2016 among others).

(1) Merge (X, Y) = {X, Y} = K

That is, Merge takes two discrete items and produces a set out of them, which itself can be an input to further Merge operation as a single item (K). Obviously, the procedure defined in this way can describe the hierarchical nature of syntax. However, what evolutionary linguists should crave is not a way of *defining* this hierarchical nature but a specific *explanation* of it. Succinctly put, a definition is not an explanation. This suggests that defining the syntactic process in this way is very harmful for evolutionary investigations of language as it sweeps one of the most important *explananda* under the rug: i.e., the hierarchical nature.

Based on this observation, I propose a decompositional approach to this procedure. Specifically, the syntactic process can be segmented into four parts: (i) Select process that maps inputs to the next process onto its workspace, (ii) Set-forming process that successively combines the materials mapped via (i), (iii) Transfer process that sends old inputs to the process (ii) to a stack, and (iv) Pushdown Stacking process that stores at most two Transferred outputs of (ii) (Chesi, 2015), making them accessible to (i). I show that all these processes cowork in a cyclic and self-organizational manner to yield a hierarchical structure; it roughly proceeds in the following fashion: (i)→(ii)→(iii)→(iv)→(i)→(ii)… The reason for the self-organization of this model is that the process (iii) makes its inputs inaccessible to the process (ii), necessarily restricting the accessible/active materials to the process (ii) (see already Boeckx, 2014). In other words, (i) gives the inputs to the process (ii), which in turn successively combines the Selected items, and (iii) makes it possible for (ii) to pay attention only to specific materials, putting all other items inside (iv). Also, the output of the series of processes (iv) →(i)→(ii) is also necessarily hierarchical since both the materials inside (iv) and the outputs of (ii) are independent/discrete chunks, neither of which inner materials are accessible to (ii). This said, my model *explains* the mechanism of the hierarchical nature of the human language syntax without defining the syntactic process in such a way that it can *capture* it. Succinctly put, my aim is to provide a principled explanation (crucially, not a definition) of the hierarchical nature of syntax and elucidate the necessary components for it.

What is more, due to its decompositionalist spirit, this model makes it possible to establish an experimentally feasible and evolutionarily plausible theory, as it provides neuroscientists, ethologists and many others with more specific questions (e.g., do animals have (iv)? how about (i)?) than before; it offers the opportunities to further narrow down the missing link(s) between "us and them." Put another way, this model fosters the comparative research more than the standard "complex-Merge" model does because the former gives empirically testable predictions while the latter shuts out the comparative/experimental feasibility, a necessary requirement for evolutionary studies.

Additionally, I claim that chimpanzees' cup-nesting observed by Greenfield (1991, 1998) sheds light on the evolution of the human language syntax. Specifically, due to the absence of the processes (iii) and (iv), their action grammar is a string-formation. In contrast, in humans, due to the expansion of the working memory along with the enlargement of the Broca's area, they acquired a giant working memory (initially had nothing to do with language) with the pushdown stack, which plays a crucial role in the generation of hierarchy (see Tomason et al., 2009 and Carruthers, 2013 for the evolution of working memory). Independently, Fujita (2009, 2014) also argues based on Greenfield's observation that this action grammar is the precursor to the syntactic computation. What my model differs from his one is that while he assumes that the process (ii) evolved into the composite operation (ii) + (iv), I dissociate these two processes since the process (iii) is necessary in order for (iv) to be active; materials cannot be stacked separately from the active chunk(s) unless they are "forgotten" by the process. As Fujita remains agnostic to this "forgetting" process, I claim that my model is conceptually more plausible than Fujita's one. And I also argue that the process (iii) is a syntactic instantiation of what Pylyshyn (2007) calls "FINST," an elementary, domain-general procedure which segments its input from others.

References

Berwick, R. C., & Chomsky, N. (2016). *Why only us?: Language and evolution*. Cambridge, MA: MIT Press.

Boeckx, C. (2014). *Elementary syntactic structures: Prospects of a feature-free syntax*. Cambridge: Cambridge University Press.

Carruthers, P. (2013). Evolution of working memory. *PNAS*, **110**(2), 10371–10378.

Chesi, C. (2015). On directionality of phrase structure building. *Journal of psycholinguistic research*, **44**, 1−50.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (2010). Some simple evo devo theses: how true might they be for language? In R. Larson, V. Deprez and H. Yamakido (eds.), *The evolution of language: Biolinguistic perspectives*, 45−62. Cambridge: Cambridge University Press.

Fitch, W. T. (2010). *Evolution of language*. Cambridge: Cambridge University Press.

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, **303**(5656), 377−380.

Fujita, K. (2009). A prospect for evolutionary adequacy: merge and the evolution and development of human language. Biolinguistics, 3 (2–3), 128–153.

Fujita, K. (2014). Recursive merge and human language evolution. In T. Roeper, & M. Speas (eds.), *Recursion: Complexity in cognition*, 243–264. New York: Springer.

Greenfield, P. M. (1991). Language, tools, and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. Behavioral and brain sciences, **14**, 531–595.

Greenfield, P. M. (1998). Language, tools, and brain revisited. Behavioral and brain sciences, **21**, 159–163.

Pylyshyn, Z. (2007). *Things and places: How the mind connects with the world*. Cambridge, MA: MIT Press.

Tomason, M. E., Race, E., Burrows, B., W.-Gabriell, S., Glover, G. H., & Gabriell, J. D. E. (2009). Development of spatial and verbal working memory capacity in the human brain. *Journal of Cognitive Neuroscience*, **21**(2), 316–332.

AESTHETIC PERCEPTION OF LINGUISTIC FEATURES AS A POTENTIAL FACTOR IN LANGUAGE CHANGE

THERESA MATZINGER*[1] and EVA SPECKER[2]

*Corresponding Author: theresa.matzinger@univie.ac.at
[1]Dept. of English & Dept. of Cognitive Biology, University of Vienna, Austria
[2] Dept. of Basic Psychological Research, University of Vienna, Austria

Aesthetic perception of visual art or music, but also of linguistic behavior such as pronunciation, word choice or prosodic patterns is common among all human cultures (Nadal & Vartanian, 2019; Rastall, 2008). This has implications for language evolution and language change: linguistic features that are perceived as aesthetically appealing are memorized easily and used frequently, and will thus be culturally transmitted to future generations of speakers (cf. Smith & Kirby, 2008). By the same rationale, less aesthetic forms will get lost over time. Thus, we suggest that aesthetic perception poses a potential constraint on language change (e.g. sound changes or lexical borrowings; Rastall, 2008).

In our exploratory study, we investigated a crucial baseline for this assumption, namely if there were indeed differences in people's aesthetic judgements of linguistic features. Specifically, we focused on the aesthetic perception of temporal rhythmic patterns in polysyllabic words. On the one hand, words might be regarded as most aesthetic if their syllables are isochronous because isochrony has a facilitatory effect on auditory processing, and people have a general propensity for regular patterns (e.g. Ravignani & Madison, 2017). On the other hand, people also perceive irregular patterns as aesthetically appealing (e.g. Westphal-Fitch & Fitch, 2013). In that case, words with deviations from isochrony might be judged as more aesthetically pleasing than purely isochronous stimuli. Especially interesting deviations are durational changes on the final syllables of words because final elements are particularly prone to durational modifications in actual speech. For example, lengthening the final syllables of phrases is a language-universal phenomenon, and might thus be linked to positive judgements (Fletcher, 2010). On the other hand, word-final (but not phrase-final) elements are frequently reduced or deleted in spoken language, possibly being linked to negative evaluations (Kohler & Rodgers, 2001; O'Brien & Fagan, 2016). Similarly, an actual sound change in the history of English might reflect a dislike of finally shortened syllables: in Middle English, words that ended in the reduced and short vowel *schwa* lost this vowel completely (Minkova, 1991).

To explore the potential link between rhythmic patterns of words and aesthetic perception, we tested 120 native German participants on their aesthetic evaluation of artificially generated trisyllabic pseudo-words. Each participant made valence ratings of 20 words that were each presented in 3 different conditions in a random order: a) with isochronous syllables, b) with the final syllable lengthened and c) with the final syllable shortened. Each participant ranked each word twice, namely once on its 'beauty' and once on its 'likability'. These concepts are highly related but still different manifestations of aesthetic appeal (Conway & Rehding, 2013). Likability refers to purely sensual pleasure and beauty requires higher executive functions (Armstrong & Detweiler, 2008; Brielmann & Pelli, 2017). We divided our participants into two groups to test how salient modifications needed to be to affect aesthetic perception. Group 1 received stimuli with final syllables lengthened/shortened by 50%, and group 2 received stimuli with final syllables lengthened/shortened by 25% of their original duration (400 ms).

We found that in group 1, people perceived isochronous and finally lengthened words as equally aesthetic but found shortened words significantly less appealing (ANOVA: beauty: $F(2, 3597) = 26.6$, $p < 0.001$; likability: $F(2, 3597) = 31.49$, $p < 0.001$). In group 2, shortening did not have a negative influence on the aesthetic perception of the words (ANOVA: beauty: $F(2, 3597) = 0.96$, $p = 0.39$; likability: $F(2, 3597) = 0.32$, $p = 0.72$). Thus, people tolerated deviations that they would normally find less appealing, if these deviations were only small (but still above the perceptual recognition threshold; Ravignani & Madison, 2017).

People's ratings may have been influenced by the intrinsic aesthetic values of the three rhythmic patterns, but also by their native language's typical and thus frequently occurring stress patterns (Bybee, 2007). There is, however, no general agreement on whether highly or less frequent items are regarded as more aesthetic (Hekkert et al., 2003). Still, speakers of German, a stress-based language, might tolerate deviations from isochrony more than speakers from syllable-based languages (Pamies, 1999). We also report results of a follow-up study addressing these issues by considering typical word stress patterns of our participants' native language and including ratings of how natural the participants find the stimuli. Overall, this study serves as a starting point to test the role of aesthetic perception of linguistic input for the cultural evolution of linguistic patterns.

References

Armstrong, T., & Detweiler-Bedell, B. (2008). Beauty as an Emotion: The Exhilarating Prospect of Mastering a Challenging World. *Review of General Psychology*, *12*(4), 305–329. https://doi.org/10.1037/a0012558

Brielmann, A. A., & Pelli, D. G. (2017). Beauty Requires Thought. *Current Biology*, *27*(10), 1506-1513.e3. https://doi.org/10.1016/j.cub.2017.04.018

Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195301571.001.0001

Conway, B. R., & Rehding, A. (2013). Neuroaesthetics and the Trouble with Beauty. *PLoS Biology*, *11*(3), 1–5. https://doi.org/10.1371/journal.pbio.1001504

Fletcher, J. (2010). The Prosody of Speech : Timing and Rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 523–602). Hoboken: Wiley-Blackwell.

Hekkert, P., Snelders, D., & Van Wieringen, P. C. W. (2003). “Most advanced, yet acceptable”: Typicality and novelty as joint predictors of aesthetic preference in industrial design. *British Journal of Psychology*, *94*(1), 111–124. https://doi.org/10.1348/000712603762842147

Kohler, K. J., & Rodgers, J. E. J. (2001). Schwa deletion in German read and spontaneous speech. *Spontaneous German Speech: Symbolic Structures and Gestural Dynamics*, 97–123. Retrieved from http://www.ipds.uni-kiel.de/kjk/pub_exx/aipuk35/kkjr.pdf

Minkova, D. (1991). *The history of final vowels in English: The sound of muting* (Vol. 4). Berlin and New York: M. de Gruyter.

Nadal, M., & Vartanian, O. (Eds.). (2019). *The Oxford Handbook of Empirical Aesthetics*. Oxford: Oxford University Press.

O’Brien, M. G., & Fagan, S. M. B. (2016). *German phonetics and phonology : theory and practice*. New Haven: Yale University Press. Retrieved from https://ubdata.univie.ac.at/AC13690024

Pamies Bertrán, A. (1999). Prosodic typology: On the dichotomy between stress-timed and syllable-timed languages. *Language Design*, *2*, 103–130.

Rastall, P. (2008). Aesthetic responses and the “cloudiness” of language: Is there an aesthetic function of language? *La Linguistique*, *44*, 103–132. https://doi.org/10.3917/ling.441.0103

Ravignani, A., & Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, *8*, 1–13. https://doi.org/10.3389/fpsyg.2017.01820

Smith, K., & Kirby, S. (2008). Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1509), 3591–3603.

https://doi.org/10.1098/rstb.2008.0145

Westphal-Fitch, G., & Fitch, W. T. (2013). Spatial Analysis of “Crazy Quilts”, a Class of Potentially Random Aesthetic Artefacts. *PLoS ONE*, *8*(9). https://doi.org/10.1371/journal.pone.0074055

THE SYMBOLIC CAPACITY AS A BASIS OF HUMAN COGNITION IN EVOLUTION AND DEVELOPMENT

LARISSA MENDOZA STRAFFON[1]
larissa.straffon@uib.no
[1]Dept. of Psychosocial Science/ SapienCE, University of Bergen, Norway

The symbolic capacity, defined as the ability to use one thing or situation to stand for another through socially constituted meaning (Rakoczy et al. 2005), is generally considered a human trait that underlies many aspects of the characteristic complex mental abilities of our species. Although researchers disagree as to precisely which faculties and conducts are to be considered typical of *H. sapiens*, most agree that modern human behaviour has symbol-making ‘at its core’ (Nowell 2010). In this paper I discuss data from archaeology, developmental and comparative psychology which indicates that, contrary to what predominant models of human cognitive evolution claim, symbolic ability emerges early in both evolution and development and should therefore be considered as a foundation of human cognition, and not as its result.

Leading models of cognitive evolution assert that the human symbolic capacity originated relatively late in evolution, over the last 50,000 years, as the result of a long process of neural development that led to the modern human brain (Coolidge and Wynn 2009; Deacon 1997; Donald 1991; Mithen 1996). However, the archaeological research of the past three decades has shown that several of the suite of traits typically associated with modern symbolic behaviour (technological innovation, art, ritual, exchange networks, etc.) appear prior to 50,000 and sometimes even earlier than 100,000 years ago, suggesting that symbolic cognition did not come about suddenly but developed gradually alongside anatomical evolution, in a stepwise cumulative process that took hundreds of thousands of years (McBrearty & Brooks 2000; Straffon 2019).

Recent data further support the presence of early symbolic cognition in extinct hominins. For example, a geometrically incised shell attributed to *H. erectus* dated c. 500,000 BP (Joordens et al. 2015), engraved bones from China made potentially by Denisovans over 100,000 years ago (Li et al. 2019), and a series of Neanderthal finds, such as a rock engraving from Gibraltar older than 39,000 years (Rodríguez-Vidal et al. 2014), and rock paintings from Spain c. 60,000 BP

(Hoffmann et al. 2018). For early *H. sapiens*, discoveries from Blombos Cave in South Africa include a cross-hatched pattern drawn with a red ochre crayon on stone c. 73,000 BP (Henshilwood et al. 2018), and a piece of reddish-brown siltstone displaying a double chevron c. 75,000 BP (Henshilwood et al. 2009). These seemingly intentional marks, recorded among four different hominin groups, indicate that the production of external symbols, likely for communication, may be a deeply-rooted behaviour in *Homo*.

The idea that symbolism arrives late is prevalent also in developmental psychology. In classic semiotics, arbitrary symbols which bear no resemblance to their referent, are considered as a higher order of signs, more complex and difficult to acquire and interpret than so-called iconic or indexical signs which have a likeness to their referent (Deacon 2006). Therefore, traditional developmental trajectories have placed the acquisition of 'proper' symbols last (Namy 2008). Yet several studies have shown that, in fact, young children do not acquire iconic signs more easily (Bohn et al. 2018), and that iconicity, not symbolicity, might be ontologically late, requiring a long period of learning and practice (Tolar et al. 2007; Saito et al. 2014). In contrast, by the first year of life, infants develop an incipient symbolicity as they learn language, and by the third year they understand symbolic content across different domains such as pretend play and drawing (Callaghan 2008; Rakoczy et al. 2005; Tomasello 2009; Vygotsky 1978). This indicates that symbolic capacity appears earlier rather than later in ontogeny and seems more cognitively available than iconic interpretation in early life.

Finally, studies in comparative psychology suggest that symbolic cognition might have even deeper phylogenetic origins. The well-known cases of Koko the gorilla, and Kanzi the bonobo, show that great apes are at least capable of successfully learning and using (though not creating) gestural, graphic, and language-based symbols (de Waal 2001; Gillespie-Lynch et al. 2011; Heimbauer et al. 2011; Matsuzawa 2009), and research into animal signalling (e.g. Vervet monkey alarm calls) alludes to symbol-based communication being widespread in nature (Ribeiro et al. 2007). So, symbolic capacity may be a trait shared with other hominins (Shea 2011) and perhaps other primates and lineages (e.g. birds, cetaceans). But whereas the perceptual aspects of symbolism might not be 'uniquely human', some of the traits required for material symbol *production* may well be (Westphal-Fitch & Fitch 2015).

The reviewed evidence shows that the emergence of symbolic capacity in both human phylogeny and ontogeny seems earlier than predicted by dominant models of cognitive evolution and development. Research into early human behaviour should then focus not on whether symbolic ability was present but on how the use and production of symbols shaped modern cognition and culture.

References

Bohn, M., Call, J., & Tomasello, M. (2019). Natural reference: A phylo-and ontogenetic perspective on the comprehension of iconic gestures and vocalizations. *Developmental science*, *22*(2), e12757.

Callaghan, T. C. (2008). The origins and development of pictorial symbol functioning. In C. Milbrath & H.M. Trautner (Eds.), *Children's understanding and production of pictures, drawing, and art* (pp. 21-32): Hogrefe.

Coolidge, F. L., & Wynn, T. (2009). *The rise of Homo sapiens: the evolution of modern thinking*: Wiley.

de Waal, F.B.M. (2001) *The Ape and the Sushi Master: Cultural Reflexions by a Primatologist:* Basic Books.

Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*: Harvard University Press.

Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain.* New York: W.W. Norton & Company.

— (2006). The aesthetic faculty. In M. Turner (Ed.), *The artful mind: Cognitive science and the riddle of human creativity* (pp. 21-53): Oxford University Press.

Gillespie-Lynch, K., Greenfield, P. M., Lyn, H., & Savage-Rumbaugh, S. (2011). The role of dialogue in the ontogeny and phylogeny of early symbol combinations: A cross-species comparison of bonobo, chimpanzee, and human learners. *First Language, 31*(4), 442-460.

Heimbauer, Lisa A., Beran, Michael J., & Owren, Michael J. (2011). A Chimpanzee Recognizes Synthetic Speech with Significantly Reduced Acoustic Cues to Phonetic Content. *Current Biology, 21*(14), 1210-1214.

Henshilwood, C. S., d'Errico, F., & Watts, I. (2009). Engraved ochres from the middle stone age levels at Blombos Cave, South Africa. *Journal of human evolution*, *57*(1), 27-47.

Henshilwood, C. S., d'Errico, F., van Niekerk, K. L., Dayet, L., Queffelec, A., & Pollarolo, L. (2018). An abstract drawing from the 73,000-year-old levels at Blombos Cave, South Africa. *Nature*, 1.

Hoffmann, D. L., Standish, C. D., García-Diez, M., Pettitt, P. B., Milton, J. A., Zilhão, Alcolea González, J.J., Cantalejo Duarte, P., Collado, H., de Balbín, R., Lorblanchet, M., Ramos Muñoz, J., Weniger, C.Chl, and Pike, A. (2018). U-Th dating of carbonate crusts reveals Neandertal origin of Iberian cave art. *Science*, *359*(6378), 912-915.

Joordens, J. C., d'Errico, F., Wesselingh, F. P., Munro, S., De Vos, J., Wallinga, J., & Mücher, H. J. (2015). Homo erectus at Trinil on Java used shells for tool production and engraving. *Nature*, *518*(7538), 228.

Li, Z., Doyon, L., Li, H., Wang, Q., Zhang, Z., Zhao, Q., & d'Errico, F. (2019). Engraved bones from the archaic hominin site of Lingjing, Henan Province. *Antiquity*, *93*(370), 886-900.

Matsuzawa, T. (2009). Symbolic representation of number in chimpanzees. *Current opinion in neurobiology, 19*(1), 92-98.

McBrearty, S., & Brooks, A. S. (2000). The revolution that wasn't: a new interpretation of the origin of modern human behavior. *Journal of Human Evolution, 39*(5), 453-563.

Mithen, S. J. (1996). *The prehistory of the mind: the cognitive origins of art, religion and science*: Thames and Hudson.

Namy, L.L. (2008), Recognition of iconicity doesn't come for free. *Developmental Science*, 11: 841-846.

Nowell, A. (2010). Defining behavioral modernity in the context of Neandertal and anatomically modern human populations. *Annual Review of Anthropology, 39*, 437-452.

Rakoczy, H., Tomasello, M., & Striano, T. (2005). How children turn objects into symbols: A cultural learning account. *Symbol use and symbol representation. New York: Erlbaum*, 67-97.

Ribeiro, S., Loula, A., de Araújo, I., Gudwin, R., & Queiroz, J. (2007). Symbols are not uniquely human. *Biosystems*, *90*(1), 263-272.

Rodríguez-Vidal, J., D’Errico, F., Pacheco, F. G., Blasco, R., Rosell, J., Jennings, R. P., Queffelec, A., Finlayson, G., Fa, D.A., Gutiérrez López, J.M., Carrión, J. S., Negro, J.J., Finlayson, S., Cáceres, L.M., Bernal, M.A., Fernández Jiménez, S. and Finlayson, C. (2014). A rock engraving made by Neanderthals in Gibraltar. *Proceedings of the National Academy of Sciences, 111*(37), 13301-13306.

Saito, A., Hayashi, M., Takeshita, H., & Matsuzawa, T. (2014). The origin of representational drawing: a comparison of human children and chimpanzees. *Child development*, *85*(6), 2232-2246.

Shea, J. J. (2011). Homo sapiens is as Homo sapiens was. *Current Anthropology, 52*(1), 1-35.

Straffon, L. M. (2019). Evolution and the Origins of Visual Art: An Archaeological Perspective. In *Handbook of Evolutionary Research in Archaeology* (pp. 407-435): Springer.

Tolar, T. D., Lederberg, A. R., Gokhale, S., & Tomasello, M. (2007). The development of the ability to recognize the meaning of iconic signs. *Journal of Deaf Studies and Deaf Education*, *13*(2), 225-240.

Tomasello, M. (2009) *The Cultural Origins of Human Cognition*: Harvard University Press.

Vygotsky, L. S. (1978). *Mind in society* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.): Cambridge, MA: Harvard University Press.

Westphal-Fitch, G., & Fitch, W. T. (2015). Towards a comparative approach to empirical aesthetics. In J.P. Huston, M. Nadal, M. Teruel, L.F. Agnati, C.J. Cela Conde, (Eds.) *Art, aesthetics, and the brain* (pp. 385-407): Oxford University Press.

STUDYING THE GENETIC BASES OF VOCAL LEARNING IN BATS

JANINE MENGEDE*[1], PAOLO DEVANNA[1], STEPHEN G. HÖRPEL[2], UWE FIRZLAFF[2], SONJA C. VERNES[1,3]

*Corresponding Author: janine.mengede@mpi.nl

[1]Neurogenetics of Vocal Communication, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

[2]Lehrstuhl für Zoologie, Technical University Munich, Munich, Germany

[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

Animal studies provide a unique opportunity to study direct causative links between genes and speech relevant behaviours, not possible in humans. Vocal learning is the ability to alter vocal outputs in response to auditory input, and is crucial for human speech (Janik & Slater, 2000). Bats are the only animals amongst the known vocal learning mammals (also including whales, dolphins, seals and elephants) that can be currently used for genetic manipulation studies, due to their small size, ease of handling and generation time (Janik & Slater, 1997; Knörnschild, 2014; Vernes, 2017).

The vast majority of studies on the neurogenetic encoding of vocal learning in the brain has been conducted in songbirds. Although this work has given us valuable insight into the genetic underpinnings of vocal learning in birds (Heston & White, 2015; Jarvis et al., 2000; Mendoza et al., 2015), their genetic and anatomical distance from humans argues for the need for similar work to be performed in mammals in order to be able to understand shared mechanisms underlying vocal learning.

FOXP2 has garnered much attention since it was identified as the first monogenetic cause of a severe speech and language disorder in humans (Lai, Fisher, Hurst, Vargha-Khadem, & Monaco, 2001). Affected individuals show structural and functional brain alterations, underlying its importance for normal brain development involved in speech and language pathways (Liegeois et al., 2003; Watkins et al., 2002). In elegant studies in songbirds, the importance of FoxP2 in vocal learning behaviour was shown by knockdown of the gene in area

X of juvenile birds, which led to disruption of song learning (Haesler et al., 2007). Taken together these data from humans and birds suggest FoxP2 as a top candidate gene to play a role in neurogenetic pathways underlying vocal learning.

Given its role in speech and language in humans and vocal learning in songbirds, we aim to determine if FoxP2 has an evolutionarily conserved role in vocal learning behaviour in bats. We have designed and tested an shRNA construct that targets the FoxP2 mRNA sequence of the pale spear-nosed bat (*Phyllostomus discolor)* and produces a functional loss of FoxP2 in cells. We are studying the effects of introducing this shRNA construct into the striatum of adult bats to determine the consequences of FoxP2 loss. In particular, histological and RNA-sequencing techniques will be used to determine the neuro-molecular consequences of FoxP2 knockdown. In the future, establishing this new knockdown model will enable demonstration of direct causative links between FoxP2 and vocal learning behaviour in a mammalian system using recently established paradigms (Lattenkamp, Vernes, & Wiegrebe, 2018). These data will demonstrate the role of FoxP2 in vocal learning mammals, bridge the evolutionary gap between birds and humans, and give a more complete picture on the evolution of vocal learning and ultimately human speech.

References

Haesler, S., Rochefort, C., Georgi, B., Licznerski, P., Osten, P., & Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS Biol, 5*(12), e321. doi: 10.1371/journal.pbio.0050321

Heston, J. B., & White, S. A. (2015). Behavior-linked FoxP2 regulation enables zebra finch vocal learning. *J Neurosci, 35*(7), 2885-2894. doi: 10.1523/jneurosci.3715-14.2015

Janik, & Slater, P. (1997). Vocal learning in mammals. *Advances in the Study of Behaviour, 26*, 59-100.

Janik, & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Anim Behav, 60*(1), 1-11. doi: 10.1006/anbe.2000.1410

Jarvis, E. D., Ribeiro, S., da Silva, M. L., Ventura, D., Vielliard, J., & Mello, C. V. (2000). Behaviourally driven gene expression reveals song nuclei in hummingbird brain. *Nature, 406*(6796), 628-632. doi: 10.1038/35020570

Knörnschild, M. (2014). Vocal production learning in bats. *Curr Opin Neurobiol, 28*, 80-85. doi: 10.1016/j.conb.2014.06.014

Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F., & Monaco, A. P. (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature, 413*(6855), 519.

Lattenkamp, E. Z., Vernes, S. C., & Wiegrebe, L. (2018). Volitional control of social vocalisations and vocal usage learning in bats. *J Exp Biol, 221*(Pt 14). doi: 10.1242/jeb.180729

Liegeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nat Neurosci, 6*(11), 1230-1237. doi: 10.1038/nn1138

Mendoza, E., Tokarev, K., During, D. N., Retamosa, E. C., Weiss, M., Arpenik, N., & Scharff, C. (2015). Differential coexpression of FoxP1, FoxP2, and FoxP4 in the Zebra Finch (Taeniopygia guttata) song system. *J Comp Neurol, 523*(9), 1318-1340. doi: 10.1002/cne.23731

Vernes, S. C. (2017). What bats have to say about speech and language. *Psychon Bull Rev, 24*(1), 111-117. doi: 10.3758/s13423-016-1060-3

Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., . . . Gadian, D. G. (2002). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain, 125*(Pt 3), 465-478. doi: 10.1093/brain/awf057

A GAME THEORETIC PERSPECTIVE ON THE UTTERANCE SELECTION MODEL FOR LANGUAGE CHANGE

JEROME MICHAUD*[1,2]

*Corresponding Author: jerome.michaud@soc.uu.se
[1]Department of Sociology, University of Uppsala, Uppsala, Sweden
[2]Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden

Many mathematical models for language change have recently been proposed and their interpretation is not always straightforward. In this paper, we take a closer look at the Utterance Selection Model with preferences and reinterpret its dynamics in terms of evolutionary game theory. The analysis demonstrates that the preference associated with a variant is formally equivalent to the payoff of using that variant in the associated game. Importantly, this payoff is subjectively perceived by speakers and evolves according to the current local use of the variant. Additionally, the accommodation to others' utterances can be encoded as a mutation term in a Replicator-Mutator equation. This analysis demonstrates how arbitrary variants can acquire fitness through usage, allowing selective processes to take place.

1. Introduction

In recent years, a number of mathematical models of language evolution have been developed (Smith, 2014; Pierrehumbert, Stonedahl, & Daland, 2014; Mitchener, 2009) and it is sometimes cumbersome to really understand how the models actually work. In this paper, we argue that reanalysing some of these models using evolutionary game theory (Weibull, 1997; Hofbauer, 1985) can shed light on the inner working of the models and provide an evolutionary interpretation of them. More specifically, we reanalyse the Utterance Selection Model (USM) (Baxter, Blythe, Croft, & McKane, 2006) and in particular its version with preferences (USMP) (Michaud, 2019) in order to demonstrate the power of the analogy with game theory.

The USM for language change models the change in the frequency of a set of competing variants due to interactions of speakers in a population. Speakers exchange utterances (biased sample of variants) to update their state or idiolect (frequency of variant usage). The different versions of the USM encode different updating rules for a speaker state. Although the USM is linguistically simplistic, (it considers a single trait that can be instantiated in a finite number of variants) it provides some fundamental insights into the emergence and change of conventions in a given population. The version with preferences (Michaud, 2019) in focus in this paper has been used to understand self-actuation of language changes as well as their S-shaped trajectories.

The main idea underlying our analysis is that the state of a speaker in the USM, which encodes the probability of an agent to produce a given variant, can be interpreted as a mixed strategy to play a game associated with the interaction, hence, the dynamics of the USM models the evolution of players' strategies. By mapping the USMP dynamics onto a Replicator-Mutator (RM) dynamics (Hofbauer, 1985; Komarova, 2004) it is possible to reinterpret the USMP in terms of evolutionary game theory, where fitness functions and mutation matrices can be defined for every individual. It turns out that any individual's preferences used in the USMP can be interpreted as speaker-dependent fitnesses of variants in the RM dynamics, whereas accommodation to incoming utterances contributes to the mutation part of the RM dynamics.

This paper is organized as follows. In section 2, we recall the definition of the USMP. In section 3, we reanalyze the USMP through the lense of evolutionary game theory. In section 4, we discuss the implication of this interpretation for the understanding of the dynamics of the USMP and language change in general.

2. The utterance selection model with preferences (USMP)

The USM for language change and its various extensions (Baxter et al., 2006; Baxter, Blythe, Croft, & McKane, 2009; Blythe & Croft, 2012; Baxter & Croft, 2016; Stadler, 2016; Stadler, Blythe, Smith, & Kirby, 2016; Michaud, 2019) model the evolution of the use of a fixed number of variants V used in a population of N speakers connected through a static network. In every pairwise interaction a speaker i and a speaker j exchange utterances $\boldsymbol{u}$[1] and update their state $\boldsymbol{x}$ and, in the USMP, their preferences $\boldsymbol{\pi}$ as illustrated in Figure 1.[2]

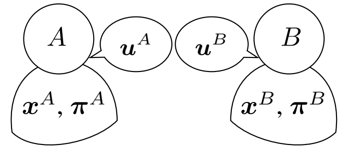


Figure 1. Illustration of the USMP interaction. Speakers A and B exchange utterances $\boldsymbol{u}$ reflecting their state $\boldsymbol{x}$ and update their state using their preferences $\boldsymbol{\pi}$.

---

[1] The utterances $\boldsymbol{u}$ are obtained from $\boldsymbol{x}$ by a biased sampling process.

[2] Bold quantities denotes column vectors of length V and are indexed to the speakers by an upper bracketed index.

The learning rule of the USMP (Michaud, 2019) is given by

$$\delta \boldsymbol{x}^{(i)} = \lambda \Big[\left(1 - h^{(ij)}\right) S\left(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(i)}\right) + h^{(ij)} A_{\mathrm{p}}\left(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(j)}, \boldsymbol{\pi}^{(i)}\right) \Big], \tag{1}$$

where λ is a learning parameter, $h^{(ij)}$ a parameter controlling the weight of the incoming utterance from speaker j with respect to speaker i own utterance, S is the self-monitoring function and A_{p} is the accommodation function with preferences. The self-monitoring and accommodation functions are given in vectorial form by[3]

$$\left\{ \begin{array}{lll} S\left(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(i)}\right) & := & \boldsymbol{u}^{(i)} - \boldsymbol{x}^{(i)}, \\ A_{\mathrm{p}}\left(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(j)}, \boldsymbol{\pi}^{(i)}\right) & := & \boldsymbol{u}^{(j)}\left(1 - \boldsymbol{\pi}^{(i)} \cdot \boldsymbol{x}^{(i)}\right) \\ & & -\mathrm{diag}\left(\boldsymbol{x}^{(i)} \otimes \left(\mathbf{1} - \boldsymbol{\pi}^{(i)}\right)\right). \end{array} \right. \tag{2}$$

The dynamics is complemented by the updating rule for the preferences

$$\delta \boldsymbol{\pi}^{(i)} = \mu(\boldsymbol{U}^{(i)} - \boldsymbol{x}^{(i)}), \qquad \boldsymbol{U}^{(i)} := \frac{1}{|V_i|} \sum_{j \in V_i} \boldsymbol{u}^{(j)}, \tag{3}$$

where μ controls the speed of change of preferences, V_i is the set of neighbors of speaker i and $\boldsymbol{U}^{(i)}$ the average uttered frequency distribution of the neighbors of a speaker i. Equations (1), (2) and (3) fully determine the dynamics of the USMP. The standard USM version is recovered when $\boldsymbol{\pi} = 0$ and $\mu = 0$ for all speakers.

The understanding of the general dynamics is fairly straightforward, at each time step, two speakers interact and their state is updated by a weighted average of their own behavior and the behavior of their interlocutor. The self-monitoring function encodes the change in the state of a speaker towards the experienced frequency of the variants. The accommodation rule is more complicated and will be the object of the game theoretic interpretation provided in this paper. In parallel with the change in the state of the speaker, the preferences are also updated and encode a social alignment process (Gaissmaier & Schooler, 2008) that occurs in addition of the interaction process; speaker i increases the preference for the variants she uses less often than her neighbors and decreases the preference for the variants she uses more often. The change in preferences has consequences on subsequent changes in the state of the speakers and enables a differential accommodation to variants.

3. Game theoretic perspective

3.1. *The USMP as a game*

Each interaction of the USMP can be concieved as a strategic interaction that can be understood in the framework of game theory. The possible variants correspond

---

[3] In Equation (2), vectors are denoted by bold symbols, a dot represents the scalar product between two vectors and the combination of the diag and the tensor product $\otimes$ is a mathematical way of encoding the elementwise multiplication of two vectors.

to the actions of the game, the state of the speakers correspond to a mixed strategy to play that game. Each time the game is played between two speakers, a variant is chosen. If the game is repeated, multiple variants can be observed with different frequencies. The utterances in the USMP, therefore, correspond to the estimated mixed strategy of the other player after L games, where L is the length of the utterance. With this interpretation, the USMP dynamics models the change in the mixed strategies of the players.

The definition of a game is incomplete without the specification of payoffs associated with the different outcome of the game. In order to extract the payoff structure associated with the USMP game, we map the accommodation rule onto a RM dynamics. By doing so, fitness functions and mutation matrices can be identified. Furthermore, using an additional assumption, one can obtain payoff matrices from the derived fitness functions.

3.2. *Link with Replicator-Mutator dynamics*

In order to explicitize the payoff matrix associated with the USMP game, we will focus on the accommodation part of the dynamics. We set $h^{(ij)} = 1$ for all i, j, scale $\lambda = \delta t$ and take the limit $\delta t \to 0$ in Equation (1). We obtain the following vectorial equation

$$\dot{\boldsymbol{x}}^{(i)} = A_{\mathrm{p}} \left(\boldsymbol{x}^{(i)}, \boldsymbol{u}^{(j)}, \boldsymbol{\pi}^{(i)} \right), \tag{4}$$

which, component-wise, reads

$$\dot{x}_v^{(i)} = u_v^{(j)} \left(1 - \sum_{w=1}^{V} \pi_w^{(i)} x_w^{(i)} \right) - x_v^{(i)} \left(1 - \pi_v^{(i)} \right). \tag{5}$$

This equation is reminiscent of the RM dynamics given by

$$\dot{x}_v = \sum_w x_w f_w(\boldsymbol{x}) Q_{wv} - \phi(\boldsymbol{x}) x_v, \quad \phi(\boldsymbol{x}) = \sum_w x_w f_w(\boldsymbol{x}), \tag{6}$$

where x_v is the fraction of the population in state v, $f(\boldsymbol{x})$ is the fitness function and Q is a mutation matrix.

In the case of the USMP game, each speaker has a different RM dynamics associated with her and we have the following correspondence

$$f_v^{(i)}(\boldsymbol{x}) \quad := \quad \pi_v^{(i)} - 1, \tag{7a}$$

$$Q_{wv}^{(ij)} \quad := \quad x_v^{(i)} - u_v^{(j)} + \delta_{vw}, \tag{7b}$$

where δ_{vw} is the Kronecker delta that equals 1 if $v = w$ and 0 otherwise, and where the fitness function and the mutation matrix are indexed to the speaker i. We will, therefore, refer to these fitness functions as *subjective*. Interestingly, we observe that the preferences define the fitness of a variant and that accommodation to incoming utterances defines the mutation matrix. Note that if the uttered variant frequency $u_v^{(j)}$ equals the expected variant frequency $x_v^{(i)}$, then there is no mutation. In general, the components of the mutation matrix reflect the fact that the larger the difference between the uttered variant frequency and the expected variant frequency, the larger the mutation rate towards the uttered variant.

From the RM dynamics, it is possible to extract an associated payoff matrix, provided that the subjective fitness functions are linear, that is we can rewrite $f(\boldsymbol{x}) = P\boldsymbol{x}$. For Equation (7), we obtain the following subjective payoff matrices

$$P^{(i)} = \begin{bmatrix} \pi_1^{(i)} - 1 & \dots & \pi_1^{(i)} - 1 \\ \vdots & \vdots & \vdots \\ \pi_V^{(i)} - 1 & \dots & \pi_V^{(i)} - 1 \end{bmatrix}. \tag{8}$$

The subjective payoff matrices $P^{(i)}$ have the property to be constant along the rows, which means that the behavior of a speaker is independent of the behavior of its interlocutor. In the USMP, there is no adaptation or strategic thinking that depends on who is talking to whom.

Under this interpretation, a rational player who tries to maximise her payoff should always choose the variant she prefers. In the USMP, the speakers play the game according to their mixed strategies, and not to the optimal rational choice, but if all speakers prefer the same variant, this variant will be used most of the time and, therefore, be conventional.

In the original version of the USM (Baxter et al., 2006), which corresponds to setting all preferences to 0, every variant is equally fit and the dynamics is purely driven by the mutation term of the RM dynamics. This explains why the original version of the USM produces more stochastic time series of variants' frequency.

Unlike conventional evolutionary games, in the USMP the payoff structure is dynamic since the preferences are updated during the dynamics. As discussed above, the change in preferences is driven by social alignment, which means that subjectively fit variants are the one used by the speech community of a speaker. Furthermore, the mutation matrices are changing at each interaction and depend on the current state of a speaker and the utterance of her interlocutor in that specific interaction.

4. Discussion

The game theoretic perspective on the USMP provided in this paper allows us to interpret the accommodation rule of this model in terms of subjective fitnesses and mutation matrices, where the preferences play the role of subjective fitnesses, while the incoming utterances drives mutations in the RM dynamics.

The RM dynamics obtained by analyzing the USMP is not standard. The fitnesses are usually negative and the mutation matrices do not satify usual properties of standard mutation matrices[4], since some of their elements can be greater than 1 or negative. This encodes the fact that a variant can mutate into itself in some kind of duplication process, while conserving the probability distribution character of the state of the speaker.

The advantage of the RM formulation is that it clearly distinguishes the role of the accommodation to the incoming utterance and the role of preferences. In the standard USM, there are no preferences and the dynamics are purely driven by the mutation part of the RM dynamics, while in the USMP the preferences act as subjective fitnesses that evolve according to local usage of variants, which grounds this model in a usage-based theory of language (Bybee, 2006). The parameter μ controlling the change of preferences is usually chosen small in order to better reflect reality. The role of this parameter is to model to speed of change in preferences. It can be thought of as a parameter controlling the importance of new utterances with respect to older ones. If new utterances have a great influence, then it is easier to change from one variant to another, whereas if the new utterances have a small influence, it will take more time for a new innovation to be adopted. It has been demonstrated by Michaud (2019) that such dynamics can explain the self-actuation of language changes.

Following the mapping from language change to ecology provided by Blythe and McKane (2007), one could make the analogy that every speaker is an island on which the variants, which are the analog to species, compete for being used. Their fitness evolves from observation of neighboring speakers/islands[5] and the mutation encodes a migration process, since conversation is the analog of migration from island to island. Following this interpretation, the idiolect of a speaker can be thought of as an ecosystem of competing variants whose fitnesses are dependent on the speaker. Furthermore, a population of speakers can be thought of as a network of ecosystems, where each linguistic ecosystem has its own dynamic fitness landscape. These linguistic ecosystems are not independent and exchange variants through conversation/migration and their fitness landscape evolves by observation of neighboring ecosystems/speakers.

In terms of evolutionary forces, the perspective given in this paper suggests that a neutral evolutionary model is not sufficient to account for language change. Such a neutral model can be used as a null model (Blythe, 2012) for language change, but in order to better account for the observed time series of change, variants should have different fitnesses. However, these fitnesses are not externally defined, they depend on the history of variants' usage. The dynamic properties and

---

[4] A standard mutation matrix has all its components positive and is row stochastic, which means that the sum of the elements of a row sum up to one.

[5] In ecology, such type of observation would be absent. This is an example where the analogy don't fully work. This is of course not a problem, since analogical reasoning has its limits and should not be applied blindly.

the fact that each speaker develops her own fitness landscape make the fitnesses sujective. This subjectivity of fitnesses allows us to reconcile the fact that variants are arbitrary (de Saussure, 1959), but their fitness evolves as they are being used.

Different versions of the USM for language change have proven useful to understand the dynamics of language change, but there are still some limitations to these models that should be addressed in future work. For instance, the USM assumes that communication is *always successful*, since every variant can be used. This is a strong assumption that should be relaxed in the future to account for failure in communication. Another limitation of this model illustrated by the subjective payoff matrices derived in this paper is that the speaker usage of the variants is *selfish*, since there is no adaptation to the identity of the interlocutor. While such an assumption is needed to keep the model tractable, it does not account for the richness of human interactions. This issue should be addressed in future development of these models. The relation to game theory outlined here suggests a path to achieve such a generalization, since game theory is the science of strategic interactions. In addition, the type of game with dynamic strategies and dynamic payoff is also a novel contribution to game theory itself, which may open new paths of cross-fertilization between language evolution and game theory.

References

Baxter, G. J., Blythe, R. A., Croft, W., & McKane, A. J. (2006). Utterance Selection Model of language change. *Physical Review E*, *73*(4), 046118.

Baxter, G. J., Blythe, R. A., Croft, W., & McKane, A. J. (2009). Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change*, *21*(02), 257–296.

Baxter, G. J., & Croft, W. (2016). Modeling language change across the lifespan: Individual trajectories in community change. *Language Variation and Change*, *28*(02), 129–173.

Blythe, R. A. (2012). Neutral Evolution: A null model for language dynamics. *ACS – Advances in Complex Systems*, *15*, 1150015.

Blythe, R. A., & Croft, W. (2012). S-curves and the mechanisms of propagation in language change. *Language*, *88*(2), 269–304.

Blythe, R. A., & McKane, A. J. (2007). Stochastic models of evolution in genetics, ecology and linguistics. *Journal of Statistical Mechanics: Theory and Experiment*, *2007*(07), P07018.

Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language*, *82*, 711–733.

de Saussure, F. (1959). *Course in general linguistics.* New York: The Philosophical Library, Inc.

Gaissmaier, W., & Schooler, L. J. (2008). The smart potential behind probability matching. *Cognition*, *109*(3), 416 - 422.

Hofbauer, J. (1985). The selection mutation equation. *Journal of mathematical* *biology*, *23*(1), 41–53.

Komarova, N. L. (2004). Replicator–mutator equation, universality property and population dynamics of learning. *Journal of Theoretical Biology*, *230*(2), 227–239.

Michaud, J. (2019). Dynamic preferences and self-actuation of changes in language dynamics. *Language Dynamics and Change*, *1*, 61–103.

Mitchener, W. G. (2009). A stochastic model of language change through social structure and prediction-driven instability. *language acquisition*, *24*, 25.

Pierrehumbert, J. B., Stonedahl, F., & Daland, R. (2014). A model of grassroots changes in linguistic systems. *arXiv preprint arXiv:1408.1985*.

Smith, A. D. (2014). Models of language evolution and change. *WIREs Cogn Sci*.

Stadler, K. (2016). *Direction and directedness in language change: An evolutrionary model of selection by trend-amplification.* Unpublished doctoral dissertation, University of Edinburgh.

Stadler, K., Blythe, R. A., Smith, K., & Kirby, S. (2016). Momentum in language change. *Language Dynamics and Change*, *6*(2), 171–198.

Weibull, J. W. (1997). *Evolutionary game theory.* MIT press.

HUMAN LANGUAGE AS A (UNIQUE) COMBINATION OF INHERITED (AND SHARED) DOMAIN-GENERAL COGNITIVE ABILITIES

MARIE MONTANT*[1,2], RAPHAELLE MALASSIS[1,3], TIPHAINE MEDAM[1] and JOEL FAGOT[1,2]

*Corresponding Author: marie.montant@univ-amu.fr

[1]Laboratoire de Psychologie Cognitive, CNRS and Aix-Marseille University, Marseille, France

[2]Institute for Language, Communication and the Brain, CNRS and Aix-Marseille University, Marseille, France

[3]School of Psychology and Neuroscience, University of Saint Andrews, Saint Andrews, United Kingdom

In cognitive sciences, human language evolution is classically studied by adopting a strictly human-centered approach: most studies in comparative psychology and comparative neurosciences are aimed at identifying one or several human-language-specific cognitive components (as described in the linguistic literature) or anatomical features (as described in the human anatomy literature) that are either missing to the other animal species, or insufficiently developed to compete with that of humans. These absent or poorly developed properties are proposed as partially responsible for the differences in complexity between human language and other animal communication systems, and are presented as potential "human-unique" features. We will discuss the relevance of human-centered approaches in the understanding of the evolution of human language and non-human communication systems, and propose an alternative that consists in studying inherited domain-general properties that we share with other species.

A current caveat in human-centered comparisons consists in assuming non-explicitly that nonhuman cognitive architectures must *resemble* human cognitive architecture, in parts or as a whole. Such an assumption could hold if: 1) human and nonhuman cognitive architectures had followed similar evolutionary paths and were adapted to comparable environmental, social and biological constraints, and 2) the cognitive architecture of each species was a construction made of independent (non-interacting) cognitive components that are not sensitive to developmental and phylogenic interactive factors. Given that every species has a unique history leading to a unique cognitive architecture, it seems like a vain enterprise to search for strictly identical components in humans and nonhuman animals. We will illustrate this proposition using syntax.

As an alternative to strictly human-centered approaches, we propose to examine the background of the human language function, namely the inherited domain-general elements of “the machinery required to master human language” (Saffran and Thiessen, 2008), that we share with other species. The underlying hypothesis we uphold here is that complex and phylogenetically recent cognitive functions, including language, are probably the result of intense re-use and re-combination of subsets of inherited anatomical, cognitive, behavioral components (Anderson, 2010). Phylogenetically close species might share with humans a combination of some (but not all) of these components, as a support for communication and/or other cognitive functions. For example, the serial organization and structuration of elements that we find in the processing of syntax might not be language-specific, but could derive from short term memory capacities that might as well be involved in the planning of complex motor sequences in humans (Koechlin and Jubault, 2006), in other primates or even in birds, including sequences of bird’s songs (Suzuki, Wheatcroft and Griesser, 2016). More generally, we question the relevance of using human syntax (or other human-centered cognitive, behavioral and anatomical features) as the norm or reference for establishing the level of complexity of non-human communication systems, in particular in species, like cetaceans, whose *Umwelten* and *Umgebung* (von Uexküll, 1956) and, as a consequence, whose cognitive architecture supposedly differ considerably from ours.

References

Anderson M., L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behavioural Brain Science*, 33, 245-266

Koechlin E, and Jubault T (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, *50*, 963-74

Saffran, J.R., and Thiessen, E.D. (2008). *Domain-general learning capacities*. In Blackwell Handbook of Language Development (Hoff, E. and Shatz, M. Eds.), pp.68 – 86. New-York, Blackwell Publishing Ltd

Suzuki, T. N., Wheatcroft, D., and Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communication*, 8. doi:10.1038/ncomms10986

Von Uexküll, J. (1956). Streifzüge durch die Umwelten von Tieren und Menschen: Ein Bilderbuch unsichtbarer Welten. (Sammlung: Verständliche Wissenschaft, Bd. 21.) Berlin: J. Springer (mit Kriszat G.).

GENETIC NETWORKS UNDER HUMAN-SPECIFIC REGULATION REVEALS CLUES ABOUT THE EVOLUTION OF THE MODERN LANGUAGE-READY BRAIN

JUAN MORIANO*[1,2], ALEJANDRO ANDIRKÓ[1,2], and CEDRIC BOECKX[1,2,3]

*Corresponding Author: jmoriano@ub.edu
[1]University of Barcelona, Barcelona, Spain
[2]Institute of Complex Systems, Barcelona, Spain
[3]ICREA

1. Introduction

Since the discovery of FOXP2 (Fisher & Vernes, 2015), the field of the genetics of language has acknowledged that to understand the genetic background of the faculty of language we have to shed light on the genetics of its cognitive subcomponents first. In addition, we now have the unprecedented opportunity to be able to work not only with data from other non-human species, but with that of our own closest extinct relatives, the Neanderthals and Denisovans (Pääbo, 2014), as well as an ever-growing corpus of genomic information from modern human populations. However, the challenge presented by this wealth of data is to structure knowledge in the form of testable hypotheses bridging genotype and phenotype.

We propose that perturbations in the background genomic networks of modern humans, in the form of various neurodevelopmental disorders, constitute a key step in understanding the complex interactions between different levels of analysis (genome, protein, cell, system). Crucially, this data is necessary to accurately inform us about our evolution as a species in general and the emergence of the faculty of language in particular. We present here an innovative approach that takes advantage of a database of high-frequency *Homo sapiens*-specific genetic variation in modern populations (Kuhlwilm & Boeckx, 2019) to experimentally test how gene expression programs key for brain development are influenced by human-specific signatures of regulation. We focus particularly on the often overlooked role of enhancers, promoters, the 5'/3' UTRome and cis-eQTLs (allele-specific expression level changes) affecting brain growth trajectories.

2. Results

In our studies we take as our starting point an extended paleogenetic dataset (Kuhlwilm & Boeckx, 2019) to evaluate how genetic regulation affects the expression of genes in the developing human brain, and pursue a multi-leveled evo-devo approach. Our first step focuses on the role of those enhancers and promoters, which determine when and where the genes are expressed, harboring human-specific single-nucleotide changes. The network of genes controlled by such regulatory regions are then evaluated at the cell level, where single-cell gene expression data analysis reveals genetic networks relevant for the generation and proliferation of progenitor cells in the developing cortex. The second step of this study assesses quantitatively the impact of human-specific single nucleotide changes on gene expression through a single-tissue eQTL analysis in 13 different brain tissues. In this analysis, we assess the effects of modern-specific variants in gene expression levels in the brain. These cis-eQTLs significantly overlap ($p = 0.0075$) with a *Homo sapiens* positive selection study. Overall, results pointing to genes implicated in neurodevelopment and clinical conditions (such as macrocephaly and microcephaly, developmental language impairment or various syndromes) stand out. The results of these two studies are complemented by the picture offered by 5'/3' untranslated regions that also underwent changes in modern human evolution.

The resulting comprehensive network we arrive at enables us to align genes implicated in neurodevelopmental disorders with distinct cognitive phenotypes in a way that candidate gene studies does not allow. This mosaic of interacting genetic nodes helps us elucidate the ontogeny of cognitive sub-systems that underlie the faculty of language, bridging genotype and phenotype, and providing candidate molecular pathways for experimental validation.

References

Fisher, S. E., & Vernes, S. C. (2015). Genetics and the Language Sciences. *Annual Review of Linguistics*, *1*(1), 289–310.

Kuhlwilm, M., & Boeckx, C. (2019). A catalog of single nucleotide changes distinguishing modern humans from archaic hominins. *Scientific Reports*, *9*(1).

Pääbo, S. (2014). The Human ConditionA Molecular Approach. *Cell*, *157*(1), 216–226.

THE SEEDS OF THE NOUN-VERB DISTINCTION IN THE MANUAL MODALITY: THE ROLE OF IMPROVISATION AND INTERACTION

YASAMIN MOTAMEDI*[1], KATHRYN MONTEMURRO[2], NATASHA ABNER[3], MOLLY FLAHERTY[4], SIMON KIRBY[5], SUSAN GOLDIN-MEADOW[2]

*Corresponding Author: y.motamedi@ucl.ac.uk
[1]University College London, UK
[2]University of Chicago, US
[3]University of Michigan, US
[4]Swarthmore College, US
[5]Centre for Language Evolution, University of Edinburgh, UK

The distinction between nouns and verbs has often been considered a fundamental feature of human language (Hockett, 1960; Hopper & Thompson, 1985). Research on emergent linguistic systems, including emergent sign languages as well as homesign systems, provides a unique window onto fundamental features of language, including the noun-verb distinction. Though they are created without a language model, research suggests that homesign systems nevertheless make this fundamental distinction between nouns and verbs (Abner et al., 2019; Goldin-Meadow et al., 2014). However, use by a community and transmission across generations also play a role and may lead to more systematic grammatical category distinctions. For example, Abner et al. (2019) analysed noun and verb signs in American Sign Language (ASL, a mature sign language) as well as Nicaraguan homesigners and 3 successive cohorts of signers of Nicaraguan Sign Language (NSL, an emerging sign language). They found that certain markers of the noun-verb distinction (sign order, sign size) were present even at the earliest stages of emergence but that other properties (repetitions, basehand) showed more systematic use in more mature systems (i.e., ASL, NSL Cohort 3). For example, the use of basehand gestures (using the non-dominant hand to represent an additional component of the verbal event) was more systematic in later stages of NSL development (and not present in ASL, indicating cross-linguistic variability).

We present a study that focuses on two processes that shape the evolution of noun-verb categories in emergent systems: i) improvisation (signal creation) and ii) interaction (signal use). In particular, we model two stages in sign language emergence: homesign, in which a new system is created, and the first generation (or cohort), in which a community of interacting users (here, former homesigners) is formed. Our study aims to further probe fundamental conceptual properties underlying the noun-verb distinction and how the coding of this distinction, present in improvised forms, might change due to conventionalization during interaction between users.

In our experiment, hearing non-signers improvised gestures to communicate a set of events, and then used those gestures either in interaction or in further isolated production. Thus, we model the distinction between continued use (as with adult homesigners) versus continued use *with* others (as when a community exists). Our events comprised video stimuli previously used by Abner et al. (2019), in which objects are used in either a typical or atypical context (e.g. taking a photo with a camera vs. digging with a camera). The typical scenarios are expected to elicit more verb-like forms, and the atypical scenarios, which highlight the object, to elicit more noun-like forms.

Participants across two studies took part either as individuals, or in pairs. Across conditions, participants first underwent an individual production stage, in which they had to improvise gestures for the events shown, without a partner. Following the first stage, participants in pairs took part in a communication round, taking turns to produce gestures for events with a partner, before completing a final round identical to the first individual production round (see figure 1). Individuals completed 3 comparable individual production rounds, without a partner.

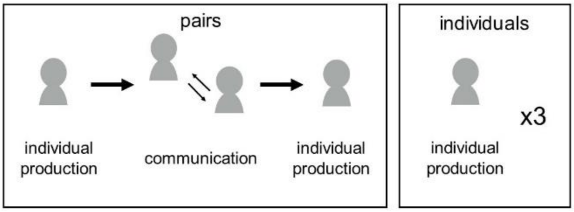


Figure 1 Stages in the experiment for pairs and individuals. Pairs produce gestures individually before communicating with each other, followed by another individual production round. Individuals complete 3 individual production rounds, with no partner present.

We analysed the parameters of gesture found to track the noun-verb distinction in sign languages by Abner et al. (2019) and other research. Preliminary results indicate similarities between the stages in our experiment and Nicaraguan homesigners and first cohort signers. Participants use both gesture order and basehand use to distinguish typical and atypical events; other parameters, such as repetitions and gesture size, do not systematically distinguish noun- and verb-like forms, but reflect iconic affordances of the events. As such, though some distinctions can be improvised, neither improvisation nor interaction by itself are sufficient for the range of formal distinctions found in naturally emerging languages.

References

Abner, N., Flaherty, M., Stangl, K., Coppola, M., Brentari, D., & Goldin-Meadow, S. (2019). The noun-verb distinction in established and emergent sign systems. *Language*, *95*(2), 230–267.

Goldin-Meadow, S., Brentari, D., Coppola, M., Horton, L., & Senghas, A. (2014). Watching language grow in the manual modality: Nominals, predicates, and handshapes. *Cognition*, *136*, 381–395.

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 88–96.

Hopper, P., J., & Thompson, S. (1985). The iconicity of the universal categories ‘noun’ and ‘verb’. In J. Haiman (Ed.), *Iconicity in Syntax*. Amsterdam: John Benjamins.

LOCALLY RESTRICTED BEHAVIORS: IN SEARCH OF POTENTIAL GREAT APE CULTURE DEPENDENT TRAITS

ALBA MOTES RODRIGO[*1] & CLAUDIO TENNIE

[*]Corresponding Author: albamotes7@gmail.com
[1]Early Prehistory and Quaternary Ecology, University of Tübingen, Tübingen, Germany

1. Introduction

Human culture is cumulative: behavioral traits are modified during the course of a generation and these modified traits are then successively copied by naïve individuals, which in turn modify the traits further (Boyd & Richerson, 1996). The main products of cumulative culture are culture dependent traits (Reindl, Apperly, Beck, & Tennie, 2017). Culture dependent traits are behavioral forms that rely on copying variants of social learning to be acquired, as due to the successive cumulation of modifications they have become too arbitrary, complex and/or opaque to be individually learned. The specifics of human languages (lexicon and gestural repertoire) are examples of culture dependent traits.

Humans are not the only species that possess culture. Other non-human animal species, including non-human great apes (henceforth apes), also have culture at least in a minimal sense (if general social learning has at least some influence on trait frequencies). However, the question remains of whether apes possess human-like culture. Whether or not apes possess cumulative culture is relevant in order to reconstruct the evolution of human culture based on the cognitive abilities of the last common ancestor (LCA) of hominins and apes.

Currently, there are two main hypotheses about the learning mechanisms underlying ape behavioral repertoires. The null hypothesis states that ape behavioral forms can be individually learned without the need to rely on copying variants of social learning (such as imitation; Tennie, Call, & Tomasello, 2009). According to this hypothesis, ape behaviors across domains are latent solutions forming latent repertoires (Tennie et al., 2009). The alternative hypothesis states that at least some ape behavioral forms rely on copying variants of social learning to be acquired and can therefore not be individually leaned (Whiten et al., 1999; Whiten et al., 2001). According to this latter hypothesis, apes could possess culture dependent traits (and by extension have cumulative culture). In order to test these hypotheses we applied a methodology adapted from Byrne (2007) that we named the Method of Local Restriction.

2. Methodology

The Method of Local Restriction identifies locally restricted behaviors, which are present in one population (or connected population cluster) of an ape species and are performed by at least two individuals (to exclude idiosyncratic behaviors). Locally unique behaviors are a special type of locally restricted behaviors that are only present in one population of one species and absent in the other ape species. We focus on locally restricted behaviors because being only present in one population could indicate that only culturally connected individuals can acquire these forms (perhaps via copying). Applying the Method of Local Restriction we conducted an exhaustive literature search across behavioral domains (tool use, non-tool use foraging behaviors, gestures, vocalizations/sounds and environment-related behaviors) and species. As a result of this search and after consultation with experts, we have compiled a list of potential culture dependent traits in apes.

3. Results and Conclusion

We have identified the staggering low number of seven locally unique behaviors across all domains and ape species. Among these behaviors, we found one vocalization ("harmonic uuh") and a sound ("grinding") produced by Sumatran orangutans. In general, our results support the null hypothesis that the vast majority of great ape behavioral forms, including gestures and vocalizations, are present in multiple unconnected great ape populations, meaning that there were multiple independent innovators who must have individually learned these forms. We therefore conclude that great ape behavioral forms (excluding for the time being locally unique behaviors) across domains can be acquired via individual learning and consequently constitute a latent repertoire that is shared to some degree among great ape species. Although our results show that a few great ape traits might be culture dependent, we also show that cultures mainly based on copying were not present in the LCA of humans and great apes and that cumulative culture evolved later in the hominin lineage.

4. References

Boyd, R., & Richerson, P. J. (1996). Why culture is common, but cultural evolution is rare. *Proceedings-British Academy, 88*, 77-94.

Byrne, R. W. (2007). Culture in great apes: using intricate complexity in feeding skills to trace the evolutionary origin of human technical prowess. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*, 577-585.

Reindl, E., Apperly, I. A., Beck, S. R., & Tennie, C. (2017). Young children copy cumulative technological design in the absence of action information. *Sci Rep, 7*, 1788.

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of cumulative culture. *Phil. Trans. R. Soc. B, 364*, 2405-2415.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., . . . Boesch, C. (1999). Cultures in chimpanzees. *Nature, 399*, 682-685.

Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., . . . Boesch, C. (2001). Charting cultural variation in chimpanzees. *Behaviour, 138*, 1481-1516.

HOW DOES SOCIAL STRUCTURE SHAPE LANGUAGE VARIATION? A CASE STUDY OF THE KATA KOLOK LEXICON

KATIE MUDD*[1], HANNAH LUTZENBERGER[2,3], CONNIE DE VOS[2], PAULA FIKKERT[2], ONNO CRASBORN[2], BART DE BOER[1]

*Corresponding Author: katie.mudd@ai.vub.ac.be
[1]Vrije Universiteit Brussel, Brussels, Belgium
[2]Center of Language Studies, Nijmegen, Netherlands
[3]International Max Planck Research School, Nijmegen, Netherlands

Sign language emergence is an excellent source of data on how language variation is conditioned. Based on the context of sign language emergence, sign languages can be classified as Deaf community sign languages (DCSL), used by a large and dispersed group of mainly deaf individuals (Mitchell & Karchmer, 2004) or as shared sign languages (SSL), which typically emerge in tight-knit communities and are shared by deaf and hearing community members (Kisch, 2008)[1].

It has been suggested that, in small, tight-knit populations, a higher degree of variation is tolerated than in large, dispersed communities because individuals can remember others' idiolects (de Vos, 2011; Thompson et al., 2019). Confirming this, Washabaugh (1986) found more lexical variation in Providence Island Sign Language, a SSL, than in American Sign Language (ASL), a DCSL. DCSLs frequently exhibit variation influenced by schooling patterns, for instance seen in the differences between ages in British Sign Language (Stamp et al., 2014), gender in Irish Sign Language (LeMaster, 2006) and race in ASL (McCaskill et al., 2011). It remains unknown how variation is conditioned in SSLs.

The present study of Kata Kolok (KK) is one of the first in-depth studies of how sociolinguistic factors shape lexical variation in a SSL. KK emerged six generations ago in a village in Bali, Indonesia due to a high incidence of hereditary deafness (Winata et al., 2012). Over half of the village knows KK, and the majority of signers are hearing (Marsaja, 2008).

We used a picture description task of 36 stimuli to study the lexical preferences of 46 deaf and hearing KK signers. These signers were sampled by age, clan and deafness, and additional sociolinguistic features are recorded. In order to study variation in the KK lexicon, we use the following three steps: 1) we focus on the first variant in the sequence produced by participants, 2) we classi-

---

[1] However, this two-category dichotomy has been questioned, given that features characterizing SSLs and DCSLs (e.g. the proportion of deafness) are continuous (Nyst, 2012).

fy signs on the basis of *underlying iconic motivation and mapping,* and 3) we compare individual repertoires by calculating the lexical distance between participants.

In line with previous findings from SSLs we find a large amount of variation in the KK lexicon. However, it appears that there is less variation for everyday concepts, e.g. *dog*, and more variation for less frequently occurring concepts, e.g. *dragonfruit.* To understand how this variation is conditioned, we compare the first signs produced by participants by creating a lexical distance matrix. We visualize this matrix using multi-dimensional scaling (MDS) (Fig. 1). Participants with a short distance from one another have similar lexical repertoires.

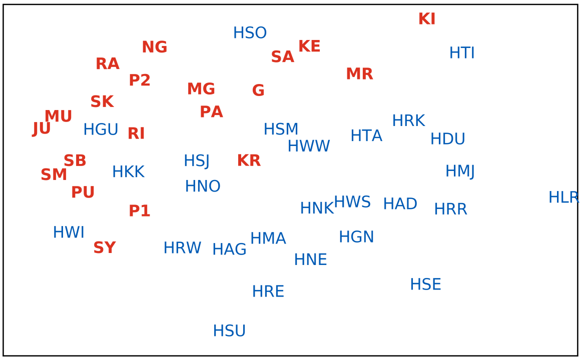


Figure 1. MDS visualization of the lexical distance between participants, with deaf participants colored in red and bolded, and hearing participants colored in blue.

Clearly, deaf and hearing participants have different lexical preferences (Fig. 1). To test this hypothesis, we use a Chi-squared test to compare two models, one with the coordinates as predictors and one without. The model with the coordinates from the MDS visualization is significantly better at predicting if participants are deaf or hearing (χ^2=12.05, df=1, p=0.00). We find that no other sociolinguistic factors significantly predict lexical variation in KK, likely due to the tight-knit community.

To conclude, in line with other SSLs, KK exhibits a high degree of variation in the lexicon, which is conditioned by deafness. Both of these results are in contrast to what has been reported for DCSLs. We conclude that the social setting of sign language emergence and evolution directly affects the degree of variation and how variation is conditioned.

References

de Vos, C. (2011). Kata Kolok color terms and the emergence of lexical signs in rural signing communities. *The Senses and Society, 6*(1), 68-76.

Kisch, S. (2008). “Deaf discourse”: the social construction of deafness in a Bedouin community. *Medical Anthropology, 27*(3), 283-313.

LeMaster, B. (2006). Language contraction, revitalization, and Irish women. *Journal of Linguistic Anthropology, 16*(2), 211-228.

Marsaja, I. G. (2008). *Desa Kolok: A deaf village and its sign language in Bali, Indonesia*. Ishara Press.

McCaskill, C., Lucas, C., Bayley, R., & Hill, J. (2011). *The hidden treasure of Black ASL: Its history and structure.* Washington, DC: Gallaudet University Press.

Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies, 4*(2), 138–163.

Nyst, V. (2012). Shared sign languages. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook.* (pp. 552–574). Berlin: Mouton de Gruyter.

Stamp, R., Schembri, A., Fenlon, J., Rentelis, R., Woll, B., & Cormier, K. (2014). Lexical Variation and Change in British Sign Language. *PLoS ONE 9*(4): e94053.

Thompson, B., Raviv, L. & Kirby, S. (2019). *Complexity can be maintained in small populations: a model of lexical variability in emerging sign languages*. Presentation at the workshop Interaction and the Evolution of Linguistic Complexity, Edinburgh.

Washabaugh, W. (1986). *Five fingers for survival*. Karoma.

Winata, S., Arhya, I. N., Moeljopawiro, S., Hinnant, J. T., Liang, Y., Friedman, T. B., & Asher, J. H. (1995). Congenital non-syndromal autosomal recessive deafness in Bengkala, an isolated Balinese village. *Journal of medical genetics*, *32*(5), 336-343.

A CULTURAL-EVOLUTIONARY MODEL OF POLITENESS STRATEGIES AND REPUTATION

ROLAND MÜHLENBERND*[1,2], SŁAWOMIR WACEWICZ[1], and PRZEMYSŁAW ŻYWICZYŃSKI[1]

*Corresponding Author: muehlenbernd@web.de
[1]Department of English Studies, Nicolaus Copernicus University, Toruń, Poland
[2]Department of Management, Ca' Foscari University, Venice, Italy

1. Introduction

Why do we say 'I am sorry' or 'thank you'? Although politeness is part and parcel of the fabric of human interaction, and as such of great interest to social psychology, human ethology and human behavioral ecology (Brown, 2015; Wacewicz, Żywiczyński, & McCrohon, 2015), research into the social costs and benefits of speaking politely is in its infancy. From a utility-based and cost-benefit perspective, many forms of polite language can be seen as instances of "strategic use of conversational politeness" (SCP). Existing utility-based accounts of SCP (cf. Pinker, 2007; Clark, 2012; Quinley, 2012) assume the costs of polite communications to be cashed out in somewhat nebulous social commodity - 'face' (Goffman, 1967; Brown & Levinson, 1987). In contrast, the recently proposed Responsibility Exchange Theory (RET) by Chaudhry and Loewenstein (2019) explains SCP in situations involving a transfer of credit or blame by grounding costs and benefits in more tangible social constructs: (perceived) competence and (perceived) warmth. In this research, we extend their epistemic model into a broader, evolutionary model of SCP.

2. Responsibility Exchange Theory

The crux of RET is in the two proposed valued social commodities, perceived competence (c) and perceived warmth (w), which jointly form the social image of a person and underwrite the communicators' utility functions. Giving away credit through *thanking* (or *apologizing*) decreases the speaker's c value, but at the same time increases of the hearer's c value. However, these two polite communications also act to increase the speaker's w value. Conversely, claiming credit through *bragging* (or *blaming*) acts exactly the other way around: it increases the speaker's c value and decreases that of the hearer, while these two rather impolite communications decrease the speaker's w value (see Table 1).

Table 1. Speaker acts, its Politeness classification, and its effect on the speaker's and hearer's social images c and w (adapted, Chaudhry & Loewenstein, 2019).

Speaker's act	Classification	Effect on speaker	Effect on Hearer
Thanking, apologizing	Polite	$-c$, $+w$	$+c$
Bragging, blaming	Impolite	$+c$, $-w$	$-c$

3. An evolutionary dynamic in a population of agents

Based on RET, we built a game-theoretic model of two agents whose behaviors affect their social image, so that in each situation an agent can gain or lose an amount of c or w depending on her strategy and that of the other agent. We develop a symmetric game where agents choose among four possible strategies that represent combinations of the speaker's acts: i) being polite P (always thank/apologize, never brag/blame), ii) being impolite IP (never thank/apologize, always brag/blame), iii) always quiet AQ (never thank/apologize/brag/blame), and iv) always communicative AC (always thank/apologize/brag/blame). As rational players, agents aim at maximizing their image-based utilities c and w.

In the next step, we used tools from evolutionary game theory (EGT) to study the dynamics and stability aspects of the game, and found the following:

1. When $c < w$ (social image in warmth is more valued than social image in competence), then the polite strategy P is the only *evolutionarily stable strategy* (Maynard Smith & Price, 1973) of the game.

2. When $c > w$ (social image in warmth is less valued than social image in competence), then none of the four strategies is evolutionarily stable and evolutionary dynamics (e.g. the replicator dynamics; Taylor & Jonker, 1978) produce a 'Cycle of Politeness', where one strategy replaces another one over time in the following order: $AQ \rightarrow IP \rightarrow AC \rightarrow P \rightarrow AQ \rightarrow ...$

4. Discussion: Politeness and reputation

Our cultural-evolutionary model of linguistic politeness describes the fitness of differently polite behavioral strategies – but can this extend to the biological fitness of their human vectors, i.e. language users? Departing from the proximate-level currencies of w and c, we propose to consider a more speculative but more ultimate-level reformulation. Accordingly, SCP often involves an interplay between the two most basic aspects of one's reputation, where perceived competence relates to a person's *status* (reputation for the capacity for acquiring and holding resources) and perceived warmth relates to one's *generosity* (reputation for the propensity for sharing resources). The important common denominator is that both acquiring and sharing resources fundamentally determine one's desirability as a cooperative partner, so the grounding of reputational payoffs invites making contact with general theories of cooperation, such as 'the leading eight' (Ohtsuki & Iwasa, 2006) or 'biological markets theory' (Noë & Hammerstein, 1995).

Acknowledgements

Roland Mühlenbernd has received funding from the European Union's Horizon 2020 research and innovation 604 programme under grant agreement no. 732942, as well as from the Polish National Agency for Academic Exchange (NAWA) under grant agreement PPN/ULM/2019/1/00222.

References

Brown, P. (2015). Politeness and language. In *International encyclopedia of the social and behavioral sciences (second edition)* (pp. 326–330). ScienceDirect.

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language usage.* Cambridge: Cambridge University Press.

Chaudhry, S., & Loewenstein, G. (2019). Thanking, apologizing, bragging, and blaming: Responsibility exchange theory and the currency of communication. *Psychological Review*, *126*(3), 313–344.

Clark, R. (2012). *Meaningful games: Exploring language with game theory.* Cambridge: MIT Press.

Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior.* New York: Doubleday.

Maynard Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*(5427), 15–18.

Noë, R., & Hammerstein, P. (1995). Biological markets. *Trends in Ecology and Evolution*, *10*(8), 336–339.

Ohtsuki, H., & Iwasa, Y. (2006). The leading eight: social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, *239*(4), 435–444.

Pinker, S. (2007). *The stuff of thought : Language as a window into human nature.* New York: Viking.

Quinley, J. (2012). Trust games as a modal request. In D. Lassiter & M. Slavkovik (Eds.), *New directions in logic, language and computation - lecture notes in computer science* (Vol. 7415, pp. 221–233). Heidelberg/New York: Springer.

Taylor, P., & Jonker, L. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, *40*, 145–156.

Wacewicz, S., Żywiczyński, P., & McCrohon, L. (2015). Linguistic politeness from an ethological perspective: Theoretical questions and empirical issues. *Theoria et Historia Scientiarum*, *11*, 81–98.

STRUCTURAL SIMILARITIES BETWEEN NATURAL LANGUAGES AND ARTIFICIAL LANGUAGES IN THE DOMAIN OF COLOR AND THEIR IMPACT ON COMMUNICATION: A LARGE-SCALE ONLINE STUDY

THOMAS F. MÜLLER*[1], JAMES WINTERS[1], TIFFANY MORISSEAU[1], IRA NOVECK[2], and OLIVIER MORIN[1]

*Corresponding Author: tmueller@shh.mpg.de
[1]Minds and Traditions Research Group, Max Planck Institute SHH, Jena, Germany
[2]Institut des Sciences Cognitives – Marc Jeannerod, CNRS, Lyon, France

Artificial language studies have well established that structure in language can arise through a tradeoff between compressibility and the functional need for expressivity (Kirby, Tamariz, Cornish, & Smith, 2015). This structure is what allows language to reduce continuous meaning spaces to discrete signals (Carr, Smith, Cornish, & Kirby, 2017), which in turn enables interlocutors to communicate successfully. One issue that has received little attention here is about the possible similarities between the structure evolving in the artificial language and natural language (even though the potential problem has been recognized early on: cf. Kirby, Cornish, & Smith, 2008). Given that participants are already fluent in at least one native language before they use the artificial one, the extent of potential biases remains unclear (cf. Xu, Dowman, & Griffiths, 2013). In this study, we focus on the domain of color, which is an example of a continuous meaning space that has been of major interest for scholars of language in the past (e.g. Berlin & Kay, 1969). Specifically, we ask: i) How closely does artificial language structure resemble the one for color terms in natural language? and ii) Are there measurable effects of this structure on performance and usage of the artificial language?

We address these questions by analyzing the data resulting from the 1-year runtime of an online smartphone application that was designed to create an artificial language. Access to the application was free and it was disseminated worldwide. The game asked participants to communicate a target out of an array of four colors to their partner, who then had to try and guess the correct color. For communication, participants were limited to using a set of black-and-white symbols only, which was provided for them at the start of the game. These symbols were selected such that they had no obvious and unambiguous relation to one color only (cf. Fig. 1 for all symbols).

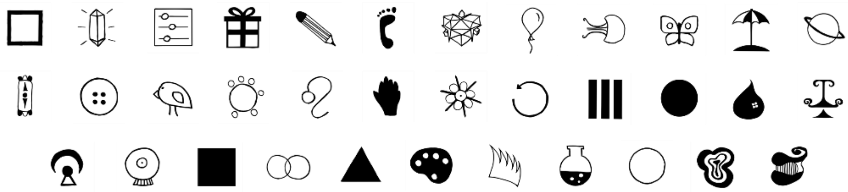


Figure 1. The 35 black-and-white symbols that players unlocked throughout the game.

Overall, we acquired data from over 4000 unique participants in this way. Crucially for our purposes, the game had been translated into 8 different languages. In this study, we need same-language participant pairs, which is why we focus on speakers of English (101 pairs), German (116), and French (44). Since we needed to ground our analyses in the structure these three languages exhibit for the 32 colors used in the game, we conducted an additional, separate study first. This was a naming task not unlike the World Color Survey (Cook, Kay, & Regier, 2005), but implemented online and with our specific color set, with 50 participants for each language.

We started our analyses by applying exploratory factor analysis to the basic color terms used in the naming task, revealing the structure in participants' native languages. The factorial structures we see (see Supplementary Material for an example) reflect the basic color terms well for English and German, but for French one unexpected factor emerged. With this baseline, we could proceed to test our predictions (pre-registered on the Open Science Framework). By applying confirmatory factor analysis, we tried to replicate the exploratory results on the artificial language data from our communication game. The results indicated at least a moderate fit for all languages, with the best outcome for German. Afterwards, we applied separate mixed-effects models to assess the effects of the individual structures found for the languages on performance and the number of symbols that participants sent. For English, we found that participants both had a harder time communicating and sent more symbols when the task presented them with colors that loaded on the same term in the factorial structure, as predicted. For German and French, no such effects could be found. These results show that the structure of artificial languages can resemble the one in natural languages, at least to some degree, and that this common structure affected both performance and usage of the artificial language, albeit only for English pairs. This implies that stimuli used in artificial languages have to be carefully tailored to participants' native languages, and that – in the bigger picture – potential biases arising from the native language should not be ignored.

References

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Univ of California Press.

Carr, J. W., Smith, K., Cornish, H., & Kirby, S. (2017). The cultural evolution of structured languages in an open-ended, continuous world. *Cognitive Science, 41*(4), 892–923.

Cook, R. S., Kay, P., & Regier, T. (2005). The World Color Survey Database. In *Handbook of Categorization in Cognitive Science* (pp. 23–241). Elsevier.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681-10686.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition, 141*, 87–102.

Xu, J., Dowman, M., & Griffiths, T. L. (2013). Cultural transmission results in convergence towards colour term universals. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1758), 20123073.

LEARNING GENERAL, ADAPTIVE AND HUMAN-INTERPRETABLE CONCEPTS THROUGH SITUATED INTERACTIONS

JENS NEVENS*[1], PAUL VAN EECKE[1], and KATRIEN BEULS[1]

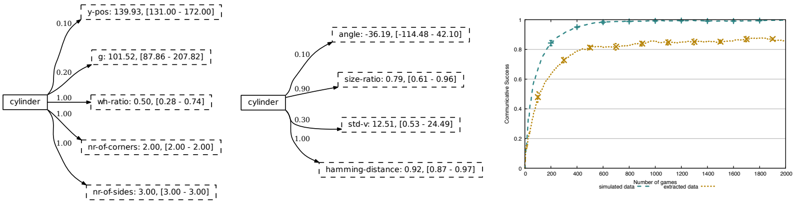
*Corresponding Author: jens@ai.vub.ac.be
[1] Artificial Intelligence Laboratory, Vrije Universiteit Brussel, Brussels, Belgium

For communicating and reasoning about the world in which we live, we rely on a repertoire of concepts that form a symbolic abstraction layer over our continuous sensori-motor experiences. For example, the cones in our eyes can convert the whole visible colour spectrum into nerve impulses, but we communicate and reason about abstractions over this spectrum, using concepts like RED, YELLOW and GREENISH BLUE. There is overwhelming evidence that these concepts are not universal or innate (Levinson, 2000), but that they can emerge and evolve through experiences and interactions in the world (Steels & Belpaeme, 2005; Bleys, 2015). The research on which we report here contributes to this view, by presenting computational simulations of how meaningful concepts can be distilled from streams of sensori-motor data through a series of situated communicative interactions. Crucially, the concepts that are learned are interpretable, adaptive to changes in the world, and general enough to be applicable to previously unseen objects.

Our approach builds further on earlier work within the language game paradigm (Steels, 2001). In this work, the concepts that were learned were either limited to continuous data on a single feature channel, such as colour (Bleys, 2015) or spatial position (Spranger, 2012), or to non-continuous data on multiple feature channels (Wellens, 2008). Here, we lift both limitations at the same time and investigate how concepts can be distilled from a larger number of continuous-valued feature channels. Our approach radically differs from other recent work, which applies deep learning techniques to concept learning, e.g. (Dolgikh, 2018; Shi, Xu, Yao, & Xu, 2019). The models resulting from these techniques are often high-performing, but require huge amounts of training data, yield concepts that are not human-interpretable, and require a partial or complete re-training of the neural network in order to adapt to changes in the world.

For our purposes, we set up a series of tutor-learner experiments, which each explore a different concept learning strategy. The experiments consist of a large number of tutor-learner interactions, which are set in a world based on the CLEVR dataset (Johnson et al., 2017). This world consists of scenes containing geometrical objects, which differ in horizontal and vertical position, colour, material, shape and size. During each interaction, the tutor uses a single word to refer to an object in the scene. The task of the learner is to point to the object to which the tutor referred. At the end of the interaction, the learner receives feedback on the outcome of the task, and the tutor points to the correct object if the learner was wrong.

The learner observes the scenes through continuous-valued human-interpretable feature channels, such as ‘area’, ‘width-height-ratio’, or ‘position-on-x-axis’. Depending on the specific experiment, the feature values are obtained through simulation or via a neural network model for object detection and segmentation (He, Gkioxari, Dollár, & Girshick, 2017; Yi et al., 2018). For each concept, the learner must simultaneously learn which feature channels are important to which extent, and what the prototypical value for each channel is. Figure 1 (a) and (b) present concepts that were learned using simulated and extracted data respectively, showing the weight and prototypical value of each relevant feature channel. Figure 1 (c) shows how the communicative success increases with the number of interactions that take place, using simulated (green line) and extracted (yellow line) features. After around 1000 interactions, a stable conceptual system is in place, achieving communicative success in 100% of the interactions using the simulated data and in nearly 85% of the interactions using the extracted data.



(a) (b) (c)

Figure 1. Learned concepts for “cylinder” using simulated (a) and extracted (b) data; communicative success over time, using simulated (green) and extracted (yellow) features (c).

In order to investigate the generality and adaptivity of the concepts that are learned, we have carried out an additional experiment. In the first phase of the experiment, the tutor and learner were placed in an environment that exhibited certain biases (condition A - e.g. cubes are always red or brown). After the conceptual system of the learner had stabilised, we then changed the biases in the world (condition B - e.g. cylinders are always blue or green). The results show no drop in communicative success when transitioning from A to B, indicating that the learned concepts are not affected by these environmental co-occurrences.

In sum, we have used computational simulations to show how human-interpretable concepts can be distilled from parallel streams of continuous sensorimotor data through repeated communicative interactions, and have demonstrated that these concepts can adapt to changing environmental conditions.

References

Bleys, J. (2015). *Language strategies for the domain of colour.* Berlin: Language Science Press.

Dolgikh, S. (2018). Spontaneous concept learning with deep autoencoder. *International Journal of Computational Intelligence Systems*, *12*(1), 1–12.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the ieee international conference on computer vision* (pp. 2961–2969). Honolulu, Hawaii.

Johnson, J., Hariharan, B., Maaten, L. van der, Fei-Fei, L., Lawrence Zitnick, C., & Girshick, R. (2017). Clevr: A diagnostic dataset for compositional language and elementary visual reasoning. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 2901–2910). Honolulu, Hawaii.

Levinson, S. C. (2000). Yélî dnye and the theory of basic color terms. *Journal of Linguistic Anthropology*, *10*(1), 3–55.

Shi, J., Xu, J., Yao, Y., & Xu, B. (2019). Concept learning through deep reinforcement learning with memory-augmented neural networks. *Neural Networks*, *110*, 47–54.

Spranger, M. (2012). Potential stages in the cultural evolution of spatial language. In *The evolution of language* (pp. 328–335). World Scientific.

Steels, L. (2001). Language games for autonomous robots. *IEEE Intelligent systems*, *16*(5), 16–22.

Steels, L., & Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and brain sciences*, *28*(4), 469–488.

Wellens, P. (2008). Coping with combinatorial uncertainty in word learning: A flexible usage-based model. In *The evolution of language* (pp. 370–377). World Scientific.

Yi, K., Wu, J., Gan, C., Torralba, A., Kohli, P., & Tenenbaum, J. (2018). Neural-symbolic vqa: Disentangling reasoning from vision and language understanding. In *Advances in neural information processing systems* (pp. 1031–1042). Montreal, Canada.

EMERGENCE OF COMMUNICATION WITH SELFISH AGENTS

MICHAEL NOUKHOVITCH*[1,2], TRAVIS LACROIX[3,2], ANGELIKI LAZARIDOU[4], and
AARON COURVILLE[1,2]

*Corresponding Author: mnoukhov@gmail.com
[1]Université de Montréal, Montréal, Canada
[2]Mila, Québec Artificial Intelligence Institute, Montréal, Canada
[3]Department of Logic & Philosophy of Science, University of California, Irvine, USA
[4]Deepmind, London, United Kingdom

Previous work on the evolution of meaning has taken advantage of the signalling-game framework—initially proposed by Lewis (1969) and extended to a dynamic setting by Skyrms (2014/1996, 2010). In the basic sender-receiver game (sometimes called a *referential* in the machine learning community), signals are used by an agent to disambiguate from a number of possible referents. There are two agents, called the sender and receiver. The sender observes some state and sends a message to the receiver; the receiver observes the message, but cannot observe the state of the world directly, and she subsequently chooses an action which determines payoffs for both agents.

A common assumption for this model is that the agents' interests are perfectly aligned, but more realistic assumptions entail that agents can have, at least partial, conflicts of interest. For example, vervet monkeys (and other mammalian species) employ simple systems of communication, but when and whether an individual sends a signal in a given context may be subject to 'audience effects'—individuals are more likely to produce alarm calls while in proximity to others than when they are alone, and they are more likely to produce alarm calls when they are in the presence of kin than when they are in the presence of non-kin (Cheney & Seyfarth, 2018). Previous work in game theory—both in the setting of economics and biology—has sought to determine whether information transfer is possible when the players have conflicting goals. For example, Skyrms (2010) considers a small number of cases where the players' interests are imperfectly aligned; Crawford and Sobel (1982) provide a more general treatment of the entailments of divergent interests; and Wagner (2012, 2014) shows that meaning can even be conveyed in a zero-sum game, though the resultant dynamics will be chaotic. When signalling is *costly*, communication may be stable—this applies equally well to economics (Spence, 1973) as it does to biology (Zahavi, 1975). In this case, a cost for sending a signal can help stabilise the possibility of honest signalling.[1] As for costless cheap talk with an existing protocol, the general take away has been that it can theoretically lead to Pareto optimal behaviour, but this is not guaranteed (Farrell & Rabin, 1996). Furthermore, these results do not take into account situations where the protocol must be learned in the first place. Many of these results depend significantly upon the game, and modelling assumptions, in question. Godfrey-Smith and Martínez (2013) analyse static signalling games to see whether common interest is a predictor of communicative viability; they show that it is possible for communication to persist in games that can be characterised as having low levels of common interest. Martínez and Godfrey-Smith (2016) complement this work with a dynamic analysis of signalling with conflict of interest using the *replicator dynamic* (Taylor & Jonker, 1978).

To study the dynamic emergence of language in situations of partial conflict, we introduce a modified sender-receiver game with a parameter that smoothly defines the game between fully cooperative (agents share a reward) and fully competitive (zero-sum). The reward of the players depends on successful communication between the sender and receiver; but, in the presence of conflict of interests, neither agent should be fully informative or trustworthy in communication. We use computer agents to play the game and train them using deep reinforcement learning to selfishly optimise their own rewards and in doing so force them to learn to communicate from scratch (Foerster, Assael, Freitas, & Whiteson, 2016; Havrylov & Titov, 2017; Lazaridou, Peysakhovich, & Baroni, 2017).

For different levels of cooperation defined by our parameter, we do a thorough search over agent configurations to try to learn effective communication with our selfish learning rule. We find that, contrary to current literature in machine learning,[2] communication always *stably* emerges between selfish agents in games that are more cooperative than competitive, without any special learning rules. In games that are more competitive, we find that basic learning rules emerge communication unstably, inefficiently, and with chaotic dynamics as the act of honest communicating is no longer a strictly dominant strategy and more resembles cooperation in a prisoner's dilemma. We test a more complex learning rule that imbues agents with a form of theory-of-mind about their opponent (Foerster et al., 2018) and find that it improves stability and efficacy of communication and allows agents to effectively learn to communicate *even* in highly competitive scenarios. Thus, we propose that communication emerges naturally in games that are more cooperative than competitive. For more competitive 2-player games, we propose three properties that are beneficial to the efficacy and stabilisation of emergent

---

[1] See also Grafen (1990b, 1990a), Maynard Smith and Harper (2003), Zollman, Bergstrom, and Huttegger (2012). Lachmann, Szamado, and Bergstrom (2001) highlight that in some cases signals need only be costly outside of equilibrium.

[2] See Cao et al. (2018), Jaques et al. (2018)

communication: when the game is iterative or played with the same opponent, when the game is general-sum, and when agents are imbued with theory-of-mind.

References

Cao, K., Lazaridou, A., Lanctot, M., Leibo, J. Z., Tuyls, K., & Clark, S. (2018). Emergent communication through negotiation. *CoRR*, *abs/1804.03980*.

Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, *15*(9), 1974–1979.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, *50*(6), 1431–1451.

Farrell, J., & Rabin, M. (1996). Cheap talk. *The Journal of Economic Perspectives*, *10*(3), 103–118.

Foerster, J. N., Assael, Y. M., Freitas, N. de, & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. In *Advances in neural information processing systems.*

Foerster, J. N., Chen, R. Y., Al-Shedivat, M., Whiteson, S., Abbeel, P., & Mordatch, I. (2018). Learning with opponent-learning awareness. In *International conference on autonomous agents and multiagent systems.*

Godfrey-Smith, P., & Martínez, M. (2013). Communication and common interest. *PLoS Computational Biology*, *9*(11), 1–6.

Grafen, A. (1990a). Biological signals as handicaps. *Journal of Theoretical Biology*, *144*, 517–546.

Grafen, A. (1990b). Sexual selection unhandicapped by the fisher process. *Journal of Theoretical Biology*, *144*, 473–516.

Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Advances in neural information processing systems* (pp. 2149–2159).

Jaques, N., Lazaridou, A., Hughes, E., Gülehre aglar, Ortega, P. A., Strouse, D., Leibo, J. Z., & Freitas, N. de. (2018). Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In *International conference on machine learning.*

Lachmann, M., Szamado, S., & Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences*, *98*(23), 13189–13194.

Lazaridou, A., Peysakhovich, A., & Baroni, M. (2017). Multi-agent cooperation and the emergence of (natural) language. In *International conference on learning representations.*

Lewis, D. (1969). *Convention: A philosophical study.* Oxford: Wiley-Blackwell.

Martínez, M., & Godfrey-Smith, P. (2016). Common interest and signaling games: a dynamic analysis. *Philosophy of Science*, *83*(3), 371–392.

Maynard Smith, J., & Harper, D. (2003). *Animal signals.* Oxford: Oxford University Press.

Skyrms, B. (2010). *Signals: Evolution, learning, & information.* Oxford: Oxford University Press.

Skyrms, B. (2014/1996). *Evolution of the social contract.* Cambridge: Cambridge University Press.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Taylor, P., & Jonker, L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, *40*, 145–156.

Wagner, E. O. (2012). Deterministic chaos and the evolution of meaning. *British Journal for the Philosophy of Science*, *63*, 547–575.

Wagner, E. O. (2014). Conventional semantic meaning in signalling games with conflicting interests. *British Journal for the Philosophy of Science*, *66*(4), 751–773.

Zahavi, A. (1975). Mate selection: A selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214.

Zollman, K. J. S., Bergstrom, C. T., & Huttegger, S. M. (2012). Between cheap and costly signals: The evolution of partially honest communication. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1750), 20121878.

ICONICITY IN SIGN GROUNDING: REPRESENTATION OR DISAMBIGUATION?

JONAS NÖLLE*[1], RICCARDO FUSAROLI[2] and KRISTIAN TYLÉN[2]

*Corresponding Author: j.nolle@sms.ed.ac.uk
[1]Centre for Language Evolution, University of Edinburgh, UK
[2]Department of Linguistics, Cognitive Science and Semiotics & Interacting Minds Centre, Aarhus University, Denmark

Recently there has been renewed interest in how communication systems ground their meaning (e.g., Harnad, 1990; Galantucci, 2005; Nölle et al., 2018), with a focus on the mechanisms underlying language evolution (Kirby, 2017). A recurrent experimental finding is that participants tend to resort to iconicity as a main strategy for bootstrapping meaning when creating a communication system from scratch (Garrod et al., 2007; Tamariz, 2017). This resonates with studies suggesting that languages are less arbitrary than often thought as they display systematic sound-meaning correspondences (Dingemanse et al., 2015; Blasi et al., 2016). While iconicity undoubtedly plays a central role in symbol grounding, we argue that a more nuanced look at grounding mechanisms beyond iconicity is overdue. We hypothesize that some of the experimental findings in laboratory studies on emergent communication might be biased by a particular task-design, where participants communicate concepts to each other using drawing, touchpad or gesture, without sharing a visual-spatial context. This leaves deictic communication (such as pointing) impossible that otherwise could point to indexicality as another fundamental grounding mechanism (Deacon, 1997; Diessel, 1999; Tomasello, 2008). We argue that even experimental observations subsumed under 'iconicity' might in fact incorporate elements of indexicality, metonymy and systematicity. In classic experimental semiotic tasks, like the Pictionary games (Garrod et al., 2007) or Silent Gesture experiments (e.g., Christensen et al., 2016; Motamedi et al., 2019) it is noticeable that discrimination in the form of identifying a target amongst competitors is central. Re-analysing data from a recent gesture study (Nölle et al., 2018) reveals how such a finite meaning space allows for strategies that build on minimal discrimination via pointing to a unique trait or using association based on common ground: E.g., in order to represent the top right stimulus in Fig. 1A, Danish participants in the study used a SMOKING gesture, since the Danish queen is known for smoking. We tested to what extent participants relied on a) this type of abstract association, that is whether gestured traits were present in the stimulus or not and b) minimally discriminating traits in the given context; in other words, the applicability of a given sign to the referential environment. We also hypothesized that informational bottlenecks from the original study could affect to what extent such semiotic strategies would be used to ground communication systems. We took the most extreme conditions, 1) the Closed+Immediate condition and 2) the Open+Displaced condition and had 3 naïve coders code how present/associative and applicable to the referential context gestured traits were on every trial (Fig 1A). We hypothesized that informational bottlenecks motivating systematicity, such as an expanding meaning space and displacement of the referential environment, should also affect how associative and generalizable gestures would become. We therefore expected more associative gestures and more gestures that did not refer to unique traits in condition 2.

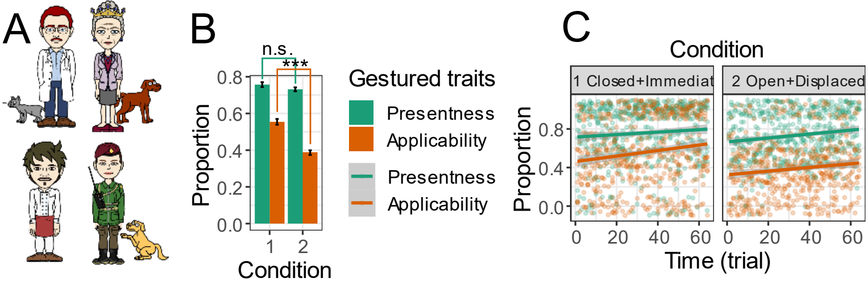


Figure 1 **A** Example Stimulus set (Note: the sets used in experiment 2 by Nölle et al., 2018 contained 16 stimuli on every trial). The referent environment determined how gestures were coded on both dimensions. E.g., the top right stimulus could be communicated with gestures expressing GLASSES (present+applicable to one target), MOUSTACHE (present+applicable to many targets), SYRINGE (associative+one), or PETTING (associative+many). **B** Proportions of presentness and applicability across all gestures produced to communicate a single stimulus. **C** Development over time.

For our analysis, we treated both variables as factors with two levels and asked for the proportion of these levels across all gestures produced on each trial. We found a main effect of applicability (p<.001) and two-way interaction between condition and factor (see Fig 1B) as well as main effect of time (p<.001, Fig 1C). There was no difference in presentness of gestured traits across conditions. See supplementary materials for full dataset and linear mixed effects models.

Our results suggest that while participants in these kind of Silent Gesture tasks prefer expressing traits that are present in the environment, a good proportion is associative (which is unaffected by informational bottlenecks). Participants are thus not only "representing" targets iconically, but grounding meaning through metonymic relations. Furthermore, pairs sharing a stable perceptual space (condition 2) seem to orient to minimally distinguishing traits more than participants who don't.

Acknowledgements

We would like to thank our coders Thea Rolskov Sloth, Signe Broedbaek and Nina Dyrberg. The study was funded by a seed grant from the Interacting Minds Centre at Aarhus University. JN was supported by a scholarship form the School of Psychology, Philosophy and Language Sciences at Edinburgh.

References

Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F., & Christiansen, M. H. (2016). Sound–meaning association biases evidenced across thousands of languages. *Proceedings of the National Academy of Sciences, 113*(39), 10818–10823.

Christensen, P., Fusaroli, R., & Tylén, K. (2016). Environmental constraints shaping constituent order in emerging communication systems: Structural iconicity, interactive alignment and conventionalization. *Cognition, 146*, 67–80.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain* (1st ed). New York: W.W. Norton.

Diessel, H. (1999). *Demonstratives: Form, function and grammaticalization* (Vol. 42). John Benjamins Publishing.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, Iconicity, and Systematicity in Language. *Trends in Cognitive Sciences, 19*(10), 603–615.

Galantucci, B. (2005). An Experimental Study of the Emergence of Human Communication Systems. *Cognitive Science, 29*(5), 737–767.

Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of Representation: Where Might Graphical Symbol Systems Come From? *Cognitive Science, 31*(6), 961–987.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena, 42*(1), 335–346.

Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin & Review, 24*(1), 118–137.

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J., & Kirby, S. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition, 192*, 103964.

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition, 181*, 93–104.

Tamariz, M. (2017). Experimental Studies on the Cultural Evolution of Language. *Annual Review of Linguistics, 3*(1), 389–407.

Tomasello, M. (2008). *Origins of human communication*. Cambridge, Mass: MIT Press.

DOES ENVIRONMENT SHAPE SPATIAL LANGUAGE? A VIRTUAL REALITY EXPERIMENT

JONAS NÖLLE*, SIMON KIRBY, JENNIFER CULBERTSON, and KENNY SMITH

*Corresponding Author: j.nolle@sms.ed.ac.uk
Centre for Language Evolution, University of Edinburgh, UK

Variation in language structure can be motivated by external factors including the social, physical and technological environment (Lupyan & Dale, 2016) and communicative pressures (Coupé, et al., 2019). An example that is still hotly debated in this regard is spatial language: Speech communities vary in the Frames of Reference (FoR) they prefer in linguistic and non-linguistic tasks (Levinson & Wilkins, 2006; e.g., egocentric *the ball is to the left of the car* vs. allocentric/geocentric *the ball is downhill of the car*), but it is unclear whether this is just due to cultural drift (Majid et al., 2004) or environmental factors (Li & Gleitman, 2002). More recently, systematic fieldwork has found that for some languages, the use of a geocentric FoR can be predicted by factors such as topography, L2 contact, education, population density and subsistence style (Bohnemeyer et al., 2015; Palmer, et al., 2017). However, it is extremely hard to disentangle such factors and their individual causal contribution (Roberts, 2018). Here we isolate topography in a controlled laboratory setting to test whether a causal relationship between spatial language and environment can be detected.

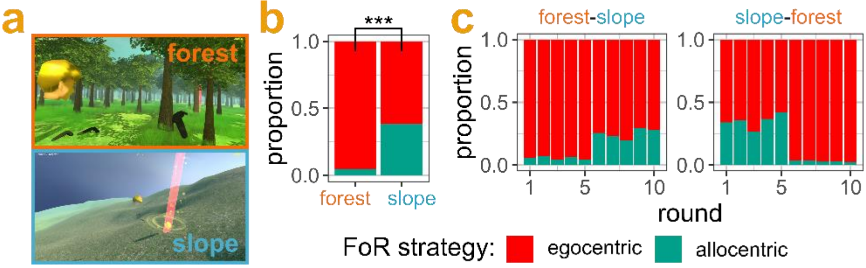


Figure 1 **a)** The two environments. **b)** Proportions of FoR strategy in both environmental conditions (Experiment 1) and **c)** over time across both experimental blocks (Experiment 2). See supplementary material for binomial linear mixed effect regressions.

For **Experiment 1** we developed *OrbHunt*, a referential spatial coordination game designed to test whether 21 dyads (n=42) speaking a predominantly egocentric language (English) would adapt their FoR depending on whether the environment in the game did or did not afford geocentric solutions. The task took place in immersive Virtual Reality (VR) to provide participants with a naïve perspective and naturalistic sense of scale; participants were placed in a forest or mountainside environment and could communicate with each other by speech. Each round a *seeker* had to collect orbs located in the environment. The seeker could only see orbs which were sufficiently close (<5m). The second player, the *director*, could see the orb from any distance but could not collect it and was invisible to the seeker, thus unable to rely on deictic devices such as pointing; therefore, the director had to describe locations using spatial language. Each dyad played 10 rounds in one of the two environmental conditions (Fig 1a), switching roles at every round. Orb collecting success was identical across conditions. While dyads used a multitude of description strategies, overall, we found that dyads on the slope relied less on the egocentric FoR and utilized more allocentric strategies, e.g., relying on the geocentric bearings *uphill*, *downhill* and *across* (Fig. 1b).

An open question is why egocentric *left/right* is so widespread and dominant (e.g., in contact situations) even though comparative phylogenetic evidence suggests it is a recent cultural innovation (Haun et al., 2006) and harder to acquire than geocentric FoR (Shusterman & Li, 2016). We hypothesize that flexibility plays a role: a *left/right* FoR strategy is useful across many environments, while specific geocentric strategies may not be. **Experiment 2** tested whether egocentric strategies are more flexible by having 20 dyads play *OrbHunt* with a change in environments after 5 rounds (switch from forest to mountainside or vice versa). However, contrary to our hypothesis, while we replicated the experiment 1 result (less egocentric FoR on the slope), we did not find a significant difference between block orders; participants readily switched to allocentric FoR even when they had played in the forest first and on the slope second (Fig. 1c).

This could be due to English conventionally allowing both FoRs, meaning there was no cost to establishing a new strategy when switching environments. We will therefore discuss a new version of experiment 2 (currently in progress) that uses an artificial communication system, where spatial descriptors must be grounded in interaction and there is a cost to switching strategies. This will show whether dyads who first negotiate an egocentric strategy fair better when presented with a new environment. In sum, we found experimental evidence for topography affecting spatial language use, which could motivate geocentric systems in the real world. Our experiments are the first to study spatial language in large-scale (rather than table-top) environments and demonstrate how VR can be used to study factors shaping language under highly controlled conditions while maintaining ecological validity. Ongoing follow-ups address potential factors explaining the rise of egocentricity, such as modern, urban mobility.

Acknowledgements

We would like to thank Hyowon Kwon, Jasmine Salmon, Elliot Huggett, Lucie Wolters and Paola Peña for their help with data collection and coding of the video material. The experimental design benefitted from discussions with Svend Østergaard and Kristian Tylén at an early conceptual stage. JN was supported by the School of Psychology, Philosophy and Language Sciences at Edinburgh.

References

Bohnemeyer, J., Donelson, K. T., Moore, R. E., Benedicto, E., Eggleston, A., O'Meara, C. K., … Gómez, M. de J. S. H. (2015). The Contact Diffusion of Linguistic Practices. *Language Dynamics and Change, 5*(2), 169–201.

Coupé, C., Oh, Y., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, *5*(9), eaaw2594.

Haun, D. B. M., Rapold, C. J., Call, J., Janzen, G., & Levinson, S. C. (2006). Cognitive cladistics and cultural override in Hominid spatial cognition. *Proceedings of the National Academy of Sciences*, *103*(46), 17568–17573.

Levinson, S. C., & Wilkins, D. P. (2006). *Grammars of Space: Explorations in Cognitive Diversity*. Cambridge University Press.

Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition*, *83*(3), 265–294.

Lupyan, G., & Dale, R. (2016). Why Are There Different Languages? The Role of Adaptation in Linguistic Diversity. *Trends in Cognitive Sciences*, *20*(9), 649–660.

Majid, A., Bowerman, M., Kita, S., Haun, D. B., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*(3), 108–114.

Palmer, B., Lum, J., Schlossberg, J., & Gaby, A. (2017). How does the environment shape spatial language? Evidence for sociotopography. *Linguistic Typology*, *21*(3), 457–491.

Roberts, S. G. (2018). Robust, Causal, and Incremental Approaches to Investigating Linguistic Adaptation. *Frontiers in Psychology*, *9*.

Shusterman, A., & Li, P. (2016). Frames of reference in spatial language acquisition. *Cognitive Psychology*, *88*, 115–161.

GLUTAMATE RECEPTORS IMPLICATED IN (SELF-)DOMESTICATION REGULATE DOPAMINERGIC SIGNALING IN STRIATAL VOCAL-LEARNING PATHWAYS

THOMAS O'ROURKE*[1,2] and CEDRIC BOECKX[1,2,3]

*Corresponding Author: thorouro10@alumnes.ub.edu
[1] Section of General Linguistics, University of Barcelona, Spain
[2] UBICS, Barcelona, Spain
[3] ICREA, Barcelona, Spain

1. Introduction: An evolutionary and neural mechanism for the emergence of complex language traits

Comparable morphological changes distinguishing anatomically modern humans (AMH) from their archaic counterparts and domesticates from their wild ancestors (e.g. brain-case shape and size alterations, retraction of the face or muzzle, and decreased tooth size) have been proposed to result from convergent evolutionary pressures (Sánchez-Villagra & Schaik, 2019). The single unifying characteristic of domesticated species, often correlating with the emergence of these physical changes, is tameness, proposed to result from an attenuation in hypothalamic-pituitary-adrenal (HPA) axis signaling, which mediates the stress response (Trut, Oskina, & Kharlamova, 2009; Wilkins, Wrangham, & Fitch, 2014; O'Rourke & Boeckx, 2019). This, in turn, raises the possibility that convergent selection, independently leading to reduced HPA-mediated stress reactivity (including reduced reactive aggression) across domesticated species and in humans, has left signals of positive selection in overlapping regions of these species' respective genomes (Wrangham, 2019; O'Rourke & Boeckx, 2019).

It has been hypothesized that the biological basis for complex language traits could have emerged as a result of a process of self-domestication in our species (Thomas & Kirby, 2018). Here, we present genomic and neurobiological evidence for how one such trait, vocal learning, may have been enhanced in modern human evolution. We present evidence that glutamatergic signaling genes — which show above-chance signals of positive selection in ours and domesticated species — are crucial regulators of the HPA axis and striatal circuits essential for vocal learning. We propose that the actions of kainate and metabotropic glutamate receptors, downregulating net excitation in stress circuits, have had concomitant modulatory effects, increasing plasticity in corticostriatal and thalamostriatal circuits crucial for vocal learning in our species.

2. Glutamate receptors in domestication and vocal learning

In a comparison of 488 neurotransmitter receptor genes across fourteen domesticated species and AMH, we have shown that glutamate receptor genes (in particular kainate and metabotropic families) show above-chance signals of positive selection, unparalleled by any other receptor type. These genes are prominently expressed in stress-response and striatal regions, and are implicated in multiple stress and striatum-related disorders, including Tourette's syndrome (O'Rourke & Boeckx, 2019; Singer, 1997; Herman, Tasker, Ziegler, & Cullinan, 2002).

Glutamate receptors are principal regulators of excitatory afferents to striatal dopaminergic circuits implicated in vocal learning. Many of the receptors we have identified function to reduce excitatory signaling acting on dopaminergic output circuits of the striatum, thus decreasing dopaminergic spiking that is often implicated in stress-induced stereotyped behaviors (O'Rourke & Boeckx, 2019; Hoffmann, Saravanan, Wood, He, & Sober, 2016; Moghaddam, 2002; Howes, McCutcheon, & Stone, 2015; Marshall, Xu, & Contractor, 2018; Xu et al., 2017).

Glutamate receptor genes are also implicated in songbird vocal-learning abilities (Wada, Sakaguchi, Jarvis, & Hagiwara, 2004). For example, the domesticated Bengalese finch, which has a reduced stress response and a more variable song repertoire than its wild vocal-learning counterpart, the white-rumped munia (Suzuki, Yamada, Kobayashi, & Okanoya, 2012; Okanoya, 2015, 2017), shows increased expression of *GRM2* in the LMAN song nucleus crucial for song variability (Okanoya, 2014). This gene shows recent signals of selection in our species (O'Rourke & Boeckx, 2019).

Other domestication and modern-human-related glutamate receptor genes (e.g. *GRM8*, *GRIK2*, and *GRIN2B*) are transcriptional targets of *FOXP2*, a gene implicated in striatally dependent vocal-learning abilities of songbirds and humans (Shi et al., 2018; Vargha-Khadem, Gadian, Copp, & Mishkin, 2005). Knockdown of *FoxP2* in the Area X song nucleus interferes with dopaminergic signaling, preventing the switch from a more variable undirected song to restricted directed singing, dependent on LMAN (Murugan, Harward, Scharff, & Mooney, 2013).

The evidence we have compiled suggests that glutamate receptor genes showing signals of positive selection in recent human evolution are implicated in reducing both stress reactivity and stereotyped vocal-learning behaviors. This raises the intriguing possibility that convergent selective pressures of (self-)domestication, attenuating the stress response in our species and domesticated songbirds, had the concomitant result of potentiating striatal-dependent vocal-learning abilities.

Acknowledgements

TOR acknowledges support from the Generalitat de Catalunya (FI 2019 fellowship). CB acknowledges support from the Spanish Ministry of Economy and Competitiveness (grant FFI2016-78034-C2-1-P/FEDER), Marie Curie International Reintegration Grant from the European Union (PIRG-GA-2009-256413), Fundaci Bosch i Gimpera, MEXT/JSPS Grant-in-Aid for Scientific Research on Innovative Areas 4903 (Evolinguistics: JP17H06379), and Generalitat de Catalunya (2017-SGR-341).

References

Herman, J. P., Tasker, J. G., Ziegler, D. R., & Cullinan, W. E. (2002). Local circuit regulation of paraventricular nucleus stress integration: Glutamate–GABA connections. *Pharmacology Biochemistry and Behavior*, *71*(3), 457–468.

Hoffmann, L. A., Saravanan, V., Wood, A. N., He, L., & Sober, S. J. (2016). Dopaminergic contributions to vocal learning. *Journal of Neuroscience*, *36*(7), 2176–2189.

Howes, O., McCutcheon, R., & Stone, J. (2015). Glutamate and dopamine in schizophrenia: an update for the 21st century. *Journal of psychopharmacology*, *29*(2), 97–115.

Marshall, J. J., Xu, J., & Contractor, A. (2018). Kainate receptors inhibit glutamate release via mobilization of endocannabinoids in striatal direct pathway spiny projection neurons. *Journal of Neuroscience*, *38*(16), 3901–3910.

Moghaddam, B. (2002). Stress activation of glutamate neurotransmission in the prefrontal cortex: implications for dopamine-associated psychiatric disorders. *Biological psychiatry*, *51*(10), 775–787.

Murugan, M., Harward, S., Scharff, C., & Mooney, R. (2013). Diminished FoxP2 levels affect dopaminergic modulation of corticostriatal signaling important to song variability. *Neuron*, *80*(6), 1464–1476.

Okanoya, K. (2014). Domestication and vocal behavior in finches. *CARTA lecture: Domestication and Human Evolution*.

Okanoya, K. (2015). Evolution of song complexity in Bengalese finches: Sexual selection and domestication as two factors. *The Journal of the Acoustical Society of America*, *138*(3), 1880–1880.

Okanoya, K. (2017). Sexual communication and domestication may give rise to the signal complexity necessary for the emergence of language: An indication from songbird studies. *Psychonomic bulletin & review*, *24*(1), 106–110.

O'Rourke, T., & Boeckx, C. (2019). Glutamate receptors in domestication and modern human evolution. *Neuroscience & Biobehavioral Reviews*.

Sánchez-Villagra, M. R., & Schaik, C. P. van. (2019). Evaluating the self-domestication hypothesis of human evolution. *Evolutionary Anthropology: Issues, News, and Reviews*.

Shi, Z., Piccus, Z., Zhang, X., Yang, H., Jarrell, H., Ding, Y., Teng, Z., Tchernichovski, O., & Li, X. (2018). miR-9 regulates basal ganglia-dependent developmental vocal learning and adult vocal performance in songbirds. *Elife*, *7*, e29087.

Singer, H. S. (1997). Neurobiology of Tourette syndrome. *Neurologic clinics*, *15*(2), 357–379.

Suzuki, K., Yamada, H., Kobayashi, T., & Okanoya, K. (2012). Decreased fecal corticosterone levels due to domestication: A comparison between the white-backed munia (lonchura striata) and its domesticated strain, the bengalese finch (lonchura striata var. domestica) with a suggestion for complex song evolution. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, *317*(9), 561–570.

Thomas, J., & Kirby, S. (2018). Self domestication and the evolution of language. *Biology & philosophy*, *33*(1-2), 9.

Trut, L., Oskina, I., & Kharlamova, A. (2009). Animal evolution during domestication: the domesticated fox as a model. *Bioessays*, *31*(3), 349–360.

Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). Foxp2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience*, *6*(2), 131.

Wada, K., Sakaguchi, H., Jarvis, E. D., & Hagiwara, M. (2004). Differential expression of glutamate receptors in avian neural pathways for learned vocalization. *Journal of Comparative Neurology*, *476*(1), 44–64.

Wilkins, A. S., Wrangham, R. W., & Fitch, W. T. (2014). The domestication syndrome in mammals: a unified explanation based on neural crest cell behavior and genetics. *Genetics*, *197*(3), 795–808.

Wrangham, R. W. (2019). *The goodness paradox: the strange relationship between virtue and violence in human evolution* (First ed.). New York: Pantheon Books.

Xu, J., Marshall, J. J., Fernandes, H. B., Nomura, T., Copits, B. A., Procissi, D., Mori, S., Wang, L., Zhu, Y., Swanson, G. T., et al.. (2017). Complete disruption of the kainate receptor gene family results in corticostriatal dysfunction in mice. *Cell reports*, *18*(8), 1848–1857.

EVOLUTION OF VOCAL COMPLEXITY: A PEAK-SHIFT AND EPIGENETICS ACCOUNT

KAZUO OKANOYA[*1],

[*]Corresponding Author: cokanoya@mail.ecc.u-tokyo.ac.jp
[1]Department of Life Sciences, The University Tokyo, Tokyo, Japan

Peak-shift is a phenomenon in which perceptual preference moves exceeding the original preference. Peak-shift occurs in two steps. First, the animal learns discrimination between two stimuli differing in one dimension. Next, when exposed with multiple stimuli including even exaggerated trait, the animal would prefer that over the original. Here we used two experiments that showed peak shift of preference for song complexity did occur in female birds and the trait to express song complexity might be inherited in male birds through epigenetic and genetic modifications. First, we trained female Bengalese finches to discriminate among different songs that differ number of trill repetitions. After training, birds preferred stimuli with longer repetitions than the original. Second, we measured the degree of methylation in one of the song control nuclei in male finches. Methylation was negatively correlated with a measure of song complexity. Together, given a set of genes that controls epigenetics of androgen receptor methylation in males, peak-shift to prefer complex songs in females would advance evolution of song complexity.

1. Introduction

Evolution of signal complexity requires two sets of account. First, the receiver animal develops preference for more complex trait than the sender animal was generally given. Second, the sender develops a tendency to enable to learn more complex signals, and this tendency should be inherited by some means.

Bengalese finches are domesticated strain of the white-rumped munia that was imported from China to Japan about 260 years ago. The bird was used as fostering parents for other species of birds because they had stronger parental behavior. Although there is no record on the selection of song complexity, domesticated Bengalese finches sing phonologically and sequentially complex songs than wild white-rumped munias (Okanoya, 2004).

One of the complexities is the number of trills in the song. Bengalese finch songs include more occurrences of repeated signals than the munias, and when it occurs the number of repetition is higher than the munias. Here we showed an evidence of behavioral peak-shift in Bengalese finch songs. We then show an evidence that song complexity and androgen methylation correlates. Together, we provide a tractable hypothesis on the evolution of song complexity.

2. Peak-shift for song complexity in Bengalese finches

Bengalese finch songs often contain trilled syllables. On average, Bengalese finch songs contain 1.9 trilled syllables in one song while it is 0.8 in munias. The number of repetitions is on average 6.7 in Bengalese and 3.6 in munias. These differences are significant, and we hypothesized the difference came about by the process of peak-shift (Terrace, 1968). We prepared two songs differing only in the trill repetition numbers (6 and 10). We trained female Bengalese finches by operant conditioning with food reinforcement to peck the response key when the song with 10 repetition of trills was played, while retain from responding when the trill repetition was 6. After discrimination learning was complete, we tested using stimuli whose trill number ranged from 5 to 11. Result showed more response to the stimulus with 11 repetitions, indicating peak shift occurred in this task (Caspani *et al*, in press).

3. Song complexity and de-methylation of a vocal control area.

In search of neural correlates of song complexity, we measured song complexity as a coefficient of variation in inter-syllable-interval in Bengalese and munias. Expression of androgen receptor correlated significantly with the song complexity. Bengalese and munias showed highly significant difference both in the level of receptor expression and the song complexity. Furthermore, androgen receptor expression was negatively correlated with methylation at the receptor coding site (Wada *et al*, 2013). Results suggest that when learning complex songs, methylation should be lower at the coding site, or alternatively, less methylation enables learning of a complex song.

4. Integrating results in into hypothesis

We showed behaviorally that the preference for trill repetition number may increase via peak-shift within the life of individual (Caspani *et al*, in press). We showed at molecular level that song complexity and level of methylation negatively correlated (Wada *et al*, in press).

Taken together, the following scenario could be drawn. Preference for song complexity develops in individual females. Learning of more complex songs in males induces less methylation at the androgen receptor coding site. Assuming that there is a set of genes responsible to modulate the tendency of methylation in Area X, and also assuming that complex song is more preferred by females, this tendency should be enhanced through generations. As a result, learning of more complex songs might became possible in Bengalese finches.

This is not at all a Lamarckian scenario of inheritance of acquired traits, but totally within the framework of Darwinian evolution. We only assumed diversity in methylation tendency that is partially heritable in the male population. We do not need to assume inheritance of preference for song complexity in females.

Acknowledgements

Supported by MEXT #4903, 17H06380.

References

Caspani, G., Fujii G. T., Mizuhara, T., & Gillman, R. T. (in press). Biased learning of sexual signals by female Bengalese Finches. *Ornithological Science*.

Okanoya, K. (2004). Song syntax in Bengalese finches: proximate and ultimate analyses. *Advances in the Study of Behavior*, 34, 297-346.

Wada, K., Hayase, S., Imai, R., Mori, C., Kobayashi, M., Liu, W. C., Takahasi, M. & Okanoya, K. (2013). Differential androgen receptor expression and DNA methylation state in striatum song nucleus Area X between wild and domesticated songbird strains. *European Journal of Neuroscience*, 38, 2600-2610.

Terrace, H. S. (1968). Discrimination learning, the peak shift, and behavioral contrast. *Journal of the Experimental Analysis of Behavior*, 11, 727-741.

LINGUISTIC EXPERIENCE AND UNIVERSAL DESIGN OF THE AUDITORY SYSTEM IN SPEECH RHYTHM PERCEPTION: IMPLICATIONS FOR SPEECH EVOLUTION

MIKHAIL ORDIN[*1,2], LEONA POLYANSKAYA[2],

DAVID GOMEZ[3], and ARTHUR G. SAMUEL[1,2,4]

[*]Corresponding Author: mikhail.ordin@¨gmail.com
[1]IKERBASQUE – Basque Foundation for Science, Bilbao, Spain
[2]BCBL, Donostia, Spain
[3]Universidad de O’Higgins, Chile
[4]Stony Brook University, New York, USA

1. Abstract

The existing literature is consistent with two plausible and reasonable hypotheses: Either (a) linguistic experience (primarily, one’s native language) shapes rhythm processing, or (b) prosody in general (and rhythmic structures in particular) in natural languages is shaped by the general design of the auditory system, cognitive mechanisms, and neural physiology. On the one hand, rhythm perception is essential for speech processing and for language acquisition in infancy (Langus et al., 2018). Rhythmic patterns differ between languages (Gervain et al., 2008; Ramus & Mehler, 1999; White & Mattys, 2007) and thus individauls may differ in their experience with different rhythms. Non-native (Polyanskaya et al., 2017; Tajima et al., 1997) or pathological (Kent et al., 1989) rhythm affects speech accentedness and comprehensibility. These observations suggest that rhythmic patterns in speech might be processed via the phonological filter of the native language. In the other hand, rhythm perception relies on a fundamental design of mammalian auditory system (Gitza, 2011; Greenberg & Ainsworth, 2004; Hickok et al., 2015; Howard & Poeppel, 2012) that underlies rhythm discrimination by animals (Tincoff et al., 2005; Toro et al., 2003) and pre-linguistic babies (Nazzi & Ramus, 2003; Ramus et al., 1999). This mechanism is not unique to humans and is shared by all people irrespective of their native language. We performed multiple experiments to

pit these two hypotheses, both logically coherent and plausible according to prior empirical evidence, against one another.

We asked participants to listen to two continuous acoustic sequences and to determine whether their rhythms were the same or different (AX discrimination). Participants were native speakers of four languages with different rhythmic properties (Spanish and French – regular rhythm; English and German – irregular rhythm), to understand whether the predominant rhythmic patterns of a native language affect sensitivity, bias and reaction time (RT) in detecting rhythmic changes in linguistic (Experiment 2) and in non-linguistic (Experiments 1 and 2) acoustic sequences. We examined sensitivity and bias measures, as well as RTs. We also computed Bayes factors in order to assess the effect of native language.

All listeners performed better (i.e., responded faster and manifested higher sensitivity and accuracy) when detecting the presence or absence of a rhythm change when the first stimulus in an AX test pair exhibited regular rhythm than when the first stimulus exhibited irregular rhythm. This result pattern was observed both on linguistic and non-linguistic stimuli and was not modulated by the native language of the participant.

We conclude that rhythmic cognition is based on general auditory and cognitive mechanisms and are not modulated by linguistic experience and are shared by all mammals. We suggest that the mechanisms are related to vocal learning, beat induction, and rhythmic entrainment (the ability to coordinate motor output with sensory input). The ability to discriminate rhythmic patterns in not only pre-requisite of speech development in ontogenesis, but also underlied speech emergence in phylogenesis of the homo genera. Irregular rhythmic patterns are marked, in a sense that any communicative system that exhibits irregular rhythm also exhibits regular rhythmic patterns, while there are communicative systems that only exhibit regular rhythms. Also, regular speech rhythms are more typologically common across languages, and the same markedness relations can be applied to speech rhythms.Regular rhythms allow better coupling between the acoustic and neural oscillations and facilitate processing of the auditory input. Thus, switching from regular to irregular rhythm can be explained by expanding the repertoir of meanings to be expressed by prosodic means and facilitate the transition to the referentiality of the signals.

References

Gervain, J. Nespor, M., Mazuka, R., Horie, R., & Mehler J. (2008) Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. Cognitive Psychology 57(1), 56-74.

Ghitza, O. (2011). Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. Front. Psychol. 2:130.doi: 10.3389/fpsyg.2011.00130

Greenberg, S. & Ainsworth, W. (2004). Speech processing in the auditory system: An Overview. In S. Greenberg, W. Ainsworth, A. Popper, & R.Fay (Eds.). Speech Processing in the Auditory System (pp. 1-62). New York: Springer Verlag.

Hickok, G., Farahbod, H., & Saberi, K. (2015). The Rhythm of Perception: Acoustic Rhythmic Entrainment Induces Subsequent Perceptual Oscillation. Psychological Science 26(7), 1006–1013.

Howard, M., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustic but not comprehension. Journal of Neurophysiology 104, 2500-2511.

Kent, R.D., Weismer, G., Kent, J.F., & Rosenbek, J.C. (1989). Toward phonetic intelligibility testing in dysarthria. Journal of Speech and Hearing Disorders 54, 482–499.

Langus, A., Mehler, J., & Nespor, M. (2018). Rhythm in language acquisition. Neuroscience & Biobehavioral Reviews 81, 158-166.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. Speech Communication 41, 233-243.

Polyanskaya, L., Ordin, M., & Busa, M. (2017). Relative Salience of Speech Rhythm and Speech Rate on Perceived Foreign Accent in a Second Language. Language and Speech 60(3), 333-355.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. Journal of the Acoustical Society of America 105 (1), 512-521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. Cognition 73(3), 265-292.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. Journal of Phonetics 25, 1-24.

Tincoff, R., Hauser, M., Tsao, F., Spaepen, G., Ramus, F., & Mehler, J. (2005). The role of speech rhythm in language discrimination: further tests with a non-human primate. Developmental Science (1), 26-35.

Toro, J.M., Trobalon, J.B, Sebastian-Galle, N (2003) The use of prosodic cues in language discrimination tasks by rats. Animal Cognition 6, 131–136.

White, L., & Mattys, S. (2007). Calibrating rhythm: First language and second language studies. Journal of Phonetics 35(4), 501-522.

COPYING FROM OBSERVING A MODEL VS. LISTENING TO LINGUISTIC INSTRUCTIONS: COMPARING CHILDREN VS. ADULTS, AND CAUSALLY RELEVANT VS. IRRELEVANT ACTIONS

ALIKI PAPA*[1], MIOARA CRISTEA[2], NICOLA McGUIGAN 3[3] and MONICA TAMARIZ[4]

*Corresponding Author: ap106@hw.ac.uk
1 School of Social Sciences, Heriot Watt University
2 School of Social Sciences, Heriot Watt University
3 School of Media, Culture and Society, University of the West of Scotland
4 School of Social Sciences, Heriot Watt University

Cultural transmission and evolution depend crucially on copying fidelity [1]. Humans tend to over-imitate: they not only copy causally-relevant actions, which contribute to achieving a goal, but also causally-irrelevant actions, which do not contribute [2-5]. We know that adding verbal input to observational, visual modelling improves copying fidelity for complex tasks [5]. However, we do not know how linguistic and visual input independently contribute to copying actions. By investigating this question, the present study suggests a possible functional factor posing selection for the evolution of language, as language would have been adaptive because it enhanced the fidelity of cultural transmission.

We investigate how transmission mode – observation (seeing someone perform an action) versus linguistic instructions (listening to how someone performed an action) – affect the probability that actions in a sequence will be copied. In addition, we compared 6 – 8-year-old children with adults, and we manipulated the actions' causal relevance.

We allocated 120 participants to 40 chains, half of children and half of adults, with 3 generations per chain. Each participant either watched silent video or listened to audio input about an action sequence leading to the extraction of a reward from a box [2]. Half of the actions were causally relevant, and half causally irrelevant. The input for each participant was the video footage or audio description of the actions in the previous participant in the chain (or the experimenter). A control group interacted with the box to determine which actions they performed spontaneously.

Apart from expected results that relevant actions are copied more than irrelevant actions, we found a significant 3-way interaction between transmission mode, relevance and age (figure 1): when it comes to causally irrelevant actions, adults overimitate in higher degrees in the Demonstration condition. However, contrary to our hypothesis, children overimitate in higher degrees in the Verbal condition. In addition, we found that both children and adults imitate more causally relevant actions in the Demonstration condition than in the Verbal condition.

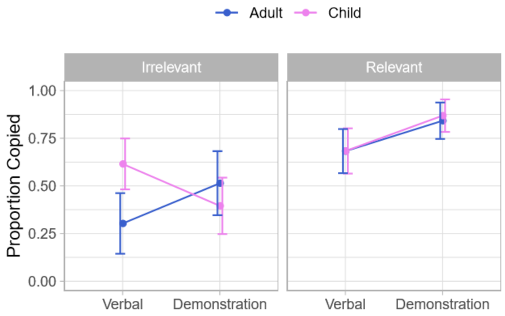


Figure 1: Here we present the 3-way interaction between transmission mode, age and relevance. Relevant actions are generally copied more in the Verbal condition, while overimitation is more common in the Demonstration condition for adults and in the Verbal condition for children..

Our results show that, amongst children, language leads to overimitation and, therefore, high fidelity transmission of behaviour. On the other hand, adults overimitate more when a model demonstrates the behaviours. We conclude that there is a point in developmental evolution of humans when language is the most efficient tool used to faithfully transmit information. After that point, however, humans learn better by observing others, i.e. in the absence of any linguistic input.

Apart from the above, we found that children mutate significantly more actions than adults do and, generally, more actions are mutated in the Demonstration condition than in the Verbal condition. Finally, there was a significant association between innovation and action relevance, with most innovations being causally relevant actions, and innovation and condition, with most innovations happening in the Verbal condition.

[1] Tennie C, Call J, Tomasello M. Ratcheting up the ratchet: on the evolution of cumulative culture. Philosophical Transactions of the Royal Society B: Biological Sciences. 2009 Aug 27;364(1528):2405-15.

[2] McGuigan N, Whiten A, Flynn E, Horner V. Imitation of causally opaque versus causally transparent tool use by 3-and 5-year-old children. Cognitive Development. 2007 Jul 1;22(3):353-64.

[3] McGuigan N, Makinson J, Whiten A. From over-imitation to super-copying: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. British Journal of Psychology. 2011 Feb 1;102(1):1-8.

[4] Nielsen M, Mushin I, Tomaselli K, Whiten A. Where culture takes hold:“Overimitation” and its flexible deployment in Western, Aboriginal, and Bushmen children. Child development. 2014 Nov;85(6):2169-84.

[5] Whiten A, Allan G, Devlin S, Kseib N, Raw N, McGuigan N. Social learning in the real-world:‘Over-imitation’occurs in both children and adults unaware of participation in an experiment and independently of social interaction. PloS one. 2016 Jul 28;11(7):e0159920.

[6] Morgan TJ, Uomini NT, Rendell LE, Chouinard-Thuly L, Street SE, Lewis HM, Cross CP, Evans C, Kearney R, de la Torre I, Whiten A. Experimental evidence for the co-evolution of hominin tool-making teaching and language. Nature communications. 2015 Jan 13;6:6029.

THE OBSTETRIC DILEMMA AND THE ORIGIN OF LANGUAGE

RICHARD PARNCUTT

parncutt@uni-graz.at
Centre for Systematic Musicology, University of Graz, Austria

Theories of language's origin can suffer from interdependent premises and conclusions. A non-circular theory might involve a biological trigger that is unique to humans and predates behavioral modernity. The obstetric dilemma is a promising candidate. Over millions of years, brain size increased and locomotion became bipedal, exacerbating parturition. Human infants became more helpless—unable to cling or crawl. Infant survival depended increasingly on complex carer-infant communication. Of the available physical signals, fundamental frequency contour (speech prosody) transmitted the most useful information in the shortest time. The human fetus became increasingly sensitive to the mother's prenatally perceptible voice, heartbeat, footstep, digestion, and movement patterns, reflecting her physical/emotional state. The fetus/infant's *mother schema* became more complex. Linguistic ability and complexity developed by statistical learning in mother/alloparent-infant interactions, building upon prenatal learning and enabled by increasing neural capacity and plasticity. The theory is consistent with different aspects of behavioral modernity including language, music, religion, and consciousness.

1. Cognitive approaches

How and why did human communication become so complex in its arbitrary sound patterns, extensive vocabularies, and hierarchical structures—relative to the poor linguistic abilities of non-human primates? The neurocognitive approach has clarified details but failed to present a complete, coherent account:

> "we do not really know how the Basic Property is actually implemented in neural circuitry. In fact … we don't have a good understanding of the range of possible implementations for any kind of cognitive computation. Our grip on how linguistic knowledge of 'grammars' might actually be implemented in the brain is even sketchier" (Berwick & Chomsky, 2017, p. 157).

Harnad (2008) asked "Where did our brains' selective capacity to learn all and only UG-compliant languages come from?" (p. 524) and emphasized the importance of avoiding circular reasoning (begging the question):

> We were looking for the evolutionary origin of the complex and abstract rules of Universal Grammar (UG). Christiansen and Chater (C&C) say … Don't ask how the UG rules evolved in the brain. The rules are in language, which is another 'organism,' not in the brain. The brain simply helped shape the language, in that the variant languages that were not learnable by the brain simply did not 'survive.' This hypothesis begs the question of why and how the brain acquired an evolved capacity to learn all and only UG-compliant languages in the first place, despite the poverty of the stimulus – which was the hard problem we started out with in the first place!

Regarding the poverty of the stimulus, data-driven statistical learning can achieve more than nativist approaches have assumed (Pullum & Scholz, 2002; Saffran, 2003). Statistical learning has also been observed in non-human primates (Hauser, Newport, & Aslin, 2001). Human infants imitate and participate in gestural-acoustic exchanges in a complex, embodied-interactive sensorimotor process (Tamis-LeMonda et al., 2014) whose fine details involve far more information than a grammatically based cognitive linguistic account.

2. The obstetric dilemma

To avoid circular reasoning, we should look for an independent biological "trigger" that catalyzed genetic and/or cultural changes leading to the emergence of complex, reflective language. The trigger should have nothing to do with language itself or indeed with any other aspect of behavioral modernity (technology, religion, art, music, consciousness) that may have interacted with language during the period when language was probably emerging (say, between 200,000 and 60,000 years ago; Berwick & Chomsky, 2017).

A promising candidate for such a trigger is the obstetric dilemma (Washburn, 1960; Wittman & Wall, 2007). During the past few million years, pre-human infants were gradually born earlier than would otherwise have been the case due to a combination of the mother's bipedal gait and the fetus's increasing brain size. The birth process (parturition) became increasingly difficult, shortening gestation, necessitating social support (midwifery), and rendering infants more helpless. While humans are not the only primates for whom birth is difficult (Leutenegger, 1974), human infants are the least able to cling or crawl. In this account of the origin of language, human infants could only survive if they developed a new kind of attachment with adults (mothers and others: Hrdy, 2009) based on acoustic-gestural communication (proto-motherese).

Approaches of this kind were introduced by Dissanayake (2000a, 2000b, 2003) and Falk (2009). Brown and Dissanayake (2018) explained:

> Such coordinated, dyadic behavior … addressed the "obstetric dilemma" of two million years ago when the anatomical trend toward a narrowed pelvis in fully bipedal *Homo erectus* conflicted at childbirth with a concomitant anatomical trend toward enlarged brains and skulls. Among other adaptations (e.g., separable pubic symphysis in females at parturition, compressible infant skull, extensive postnatal brain growth), the gestation period was significantly reduced, resulting in helpless infants dependent on their caretakers for years, rather than weeks or months as in other primates. A mother's simplification, repetition, elaboration, and exaggeration of affinitive communicative behaviors (e.g., smiling, open eyes, eyebrow flash, head bob, head nod, soft undulant vocalization, touching, patting, kissing) served to reinforce affinitive neural networks in her own brain and, when performed on a shared temporal basis, also set up a means of neural coordination of behavior and of matching of affective change between the pair.

In many species, the risk of death is highest in the first weeks, months, and years of life, which gives the events and constraints of early developmental periods special evolutionary significance. The human obstetric dilemma means that the mortal risk was even higher for early human infants due to their inability to follow or cling to a carer. The best survival strategy in this situation was to maintain the proximity and attention of mothers or alloparents by developing new forms of communication in the context of attachment (Bowlby, 1969).

3. The mother schema

The process may have involved a *mother schema* (Parncutt, 2009) that mirrors the *infant schema* of Lorenz (1943), the two interacting in motherese. Each schema is activated by specific learned or innate sensory patterns, cognitions, emotions, and behavioral interactions. The infant schema of an adult is activated by aspects of an infant's "cuteness" (visual, auditory, behavioral; multimodal). The mother schema of an infant includes its feelings toward the mother or carer and is similarly multimodal. Neither schema is confined to genetic parents or offspring, but may be generalized to other carers and dependents (cooperative breeding, allomaternal care, shared intentionality; Burkart et al., 2009; Hrdy, 2009). A child's mother schema may be activated by exposure to prenatally familiar multimodal perceptual patterns (rocking, lullabies, motherese) or by the behaviors of friendly (caring) versus dangerous (angry, careless) adults. A comparable case (multimodal, mixing "nature" and "nurture") is spontaneous attraction toward a potential sexual partner ("love at first sight")—an aspect of the *relational schemas* of Baldwin (1992).

Both motherese and crying involve vocal learning. First, there is prenatal learning of sound and movement patterns produced within the mother's body (voice, heartbeat, footsteps, digestion; Hepper, 1996; Moon & Fifer, 2000; van

Heteren et al., 2000). All such patterns are related to maternal physical and emotional state and therefore carry information that is existentially relevant for the fetus. Of the available physical signals, fundamental frequency contour (speech prosody) may transmit the most useful information in the shortest time (Coutinho & Dibben, 2013). F_0 contour represents reliable information that is physically unaffected by transmission, absorption, or reflection.

Prenatal learning can explain the extraordinary sensitivity of human newborns to maternal emotional prosody (Mastropieri & Turkewitz, 1999) and the dependence of infant crying patterns on maternal language (Mampe et al., 2009). The Pleistocene pre-human fetus/infant was more likely to survive to reproductive age if it could extract complex meanings from the prosody of its mother and other carers. Linguistic ability and complexity developed gradually between roughly 200,000 and 60,000 years ago in carer-infant interactions, upon which infant survival depended.

4. The origin of language

Simple grammatical functions may have originated as carer and infant—using different combinations of sounds, bodily gestures, and facial expressions (Pascalis et al., 2014)—labeled objects (nouns), processes (verbs), interactions (prepositions), and so on. The infant was motivated to combine these functions (cognitive "merge") because its survival in a dangerous world depended on its ability to communicate with adults. The underlying psychological mechanism was multimodal pattern recognition (Giard & Peronnet, 1999). In this way, the emergence of complex human language can be explained without invoking cognitive theories of abstract, symbolic thinking (cf. Falk, 2009).

In a statistical-learning approach, the complexity and ambiguity of human grammar and vocabulary are consistent with, and explicable by, the complexity of human neural networks (enabled by enlarged cortex) and the repetitiveness of prenatal sound patterns and infant-carer exchanges as they occur in real-world contexts—reflected by music's intrinsic repetitiveness (Margulis, 2014). The fast rate at which the human fetus and infant learn arbitrary sound-meaning relations (cf. Anderson & Thomason, 2013) kick-starts a lifelong process of cultural learning that is characteristic of humans.

The theory assumes that (proto-)motherese was more important for linguistic origins and evolution than inter-adult communication. The reason is both evolutionary and neurological: both mortality rate and neuroplasticity are higher in infancy. Today, avoidable global child mortality (mainly from disease and hunger) is about nine million per year (Black et al., 2010)—still much higher than the adult global death rate due to violence.

5. Music, religion, and consciousness

An approach to human behavioral modernity based on prenatal learning and a complex mother schema can additionally account for complex relations between sound-movement patterns and meanings in music and religion (Parncutt, 2019). It can potentially explain the extraordinarily powerful experiences that both experts and non-experts report in connection with music (Gabrielsson & Bradbury, 2011), monotheistic religion (e.g., Alston, 1993), and shamanism, including spiritual possession (Winkelman, 2004).

Musical and religious rituals may evoke the mother schema by combinations of stimuli that are familiar from prenatal life: muted light and sound, enclosed spaces, melody (reminiscent of the lowpass-filtered mother's voice), rhythm/dance (similar to the mother's heartbeat, footsteps, and associated fetal body movements), and unusual postures (fetal position). Ritual experiences that include vivid illusions of supernatural encounters are explicable if those experiences trigger participants' prenatal mother schemas.

Similarly, the soothing effect of motherese can enable carers to put down infants as they work at other tasks (Falk, 2009). The universal link between rhythm and dance (Richter & Ostovar, 2016) can be understood by considering maternal walk from the perspective of the fetus, which moves rhythmically in time with the sound of footfalls—comparable with entrainment in ensemble performance (Parncutt, 2009).

The theory can also contribute to an understanding of human reflective consciousness and its origins. Reflective consciousness involves theory of mind (Frith & Happé, 1999) and mental time-travel (imagining the past and future; Schacter et al., 2007). Both behaviors are practically unique to humans (Penn & Povinelli, 2007; Suddendorf & Corballis, 2007). The onotogeny and phylogeny of both may involve carer-infant interactions (motherese and mutual self-other consciousness; Trevarthen & Aitken, 2001), as carers and infants anthropomorphize toys and other objects, and carers creatively predict and prevent potentially fatal infant accidents.

References

Alston, W. P. (1993). *Perceiving God: The epistemology of religious experience.* Cornell University Press.

Anderson, A. L., & Thomason, M. E. (2013). Functional plasticity before the cradle: A review of neural functional imaging in the human fetus. *Neuroscience & Biobehavioral Reviews, 37*(9), 2220-2232.

Baldwin, M. W. (1992). Relational schemas and the processing of social information. *Psychological Bulletin, 112*(3), 461.

Berwick, R. C., & Chomsky, N. (2017). *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Black, R. E., Cousens, S., Johnson, H. L., Lawn, J. E., Rudan, I., Bassani, D. G., …Mathers, C. (2010). Global, regional, and national causes of child mortality in 2008: A systematic analysis. *Lancet, 375*(9730), 1969–1987.

Bowlby, J. (1969). *Attachment and loss.* New York: Basic.

Brown, S., & Dissanayake, E. (2018). The arts are more than aesthetics: Neuroaesthetics as narrow aesthetics. In M. Skov et al. (Eds.), *Neuroaesthetics* (pp. 43-57). New York: Routledge.

Burkart, J. M., Hrdy, S. B., & Van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology, 18*(5), 175-186.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*(5), 489-509.

Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition & Emotion, 27*, 658–684.

Dissanayake, E. (2000a). *Art and intimacy: How the arts began*. Seattle: University of Washington Press.

Dissanayake, E. (2000b). Antecedents of the temporal arts in early mother-infant interaction. In N. L. Wallin & B. Merker (Eds.), *The origins of music* (pp. 389-410). MIT Press.

Dissanayake, E. (2003). The core of art—Making special. *Journal of the Canadian Association for Curriculum Studies, 1*(2), 13-38.

Falk, D. (2009). *Finding our tongues: Mothers, infants, and the origins of language.* New York: Basic Books.

Frith, U., & Happé, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic? *Mind & Language, 14*(1), 82-89.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11*(5), 473-490.

Harnad, S. (2008). Why and how the problem of the evolution of Universal Grammar (UG) is hard. *Behavioral and Brain Sciences, 31*(5), 524-525.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition, 78*(3), B53–B64.

Hepper, P. G. (1996). Fetal memory: Does it exist? What does it do? *Acta Paediatrica, 416*(Supplement), 16–20.

Hrdy, S. B. (2009). *Mothers and others*. Harvard University Press.

Leutenegger, W. (1974). Functional aspects of pelvic morphology in simian primates. *Journal of Human Evolution, 3*(3), 207-222.

Lorenz, K. (1943). Die angeborenen Formen möglicher Erfahrung. *Zeitschrift für Tierpsychologie, 5,* 235-409.

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology, 19*(23), 1994-1997.

Margulis, E. H. (2014). *On repeat: How music plays the mind.* Oxford University Press.

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology, 35,* 204–214.

Moon, C. M., & Fifer, W. P. (2000). Evidence of transnatal auditory learning. *Journal of Perinatology, 20*(S1), S37.

Parncutt, R. (2009). Prenatal and infant conditioning, the mother schema, and the origins of music and religion. *Musicae Scientiae, 13*(2_suppl), 119-150.

Parncutt, R. (2019). Mother schema, obstetric dilemma, and the origin of behavioral modernity. *Behavioral Sciences, 9*(12), 142.

Pascalis, O., Loevenbruck, H., Quinn, P. C., Kandel, S., Tanaka, J. W., & Lee, K. (2014). On the links among face processing, language processing, and narrowing during development. *Child Development Perspectives, 8*(2), 65-70.

Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 731-744.

Pullum, G. K., & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *Linguistic Review, 18*(1-2), 9-50.

Richter, J., & Ostovar, R. (2016). "It don't mean a thing if it ain't got that swing"– An alternative concept for understanding the evolution of dance and music in human beings. *Frontiers in Human Neuroscience, 10,* 485.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*(4), 110-114.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience, 8*(9), 657-661.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences, 30*(3), 299-313.

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science, 23*(2), 121-126.

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(1), 3-48.

van Heteren, C. F., Boekkooi, P. F., Jongsma, H. W., & Nijhuis, J. G. (2000). Fetal learning and memory. *Lancet, 356*(9236), 1169–1170.

Washburn, S. L. (1960). Tools and human evolution. *Scientific American, 203,* 3–15.

Winkelman, M. (2004). Shamanism as the original neurotheology. *Zygon, 39*(1), 193-217.

Wittman, A. B., & Wall, L. L. (2007). The evolutionary origins of obstructed labor: bipedalism, encephalization, and the human obstetric dilemma. *Obstetrical & Gynecological Survey*, *62*(11), 739-748.

PRETEND PLAY AS A SCAFFOLD FOR LANGUAGE EVOLUTION

MICHAEL PLEYER

mpleyer@uni-koblenz.de
English Department, Universität Koblenz-Landau, Campus Koblenz, Koblenz, Germany

Pretend play is a universal human behaviour with a wide range of effects on social, cognitive, cultural, and linguistic development. Although it has deep evolutionary roots, humans have constructed a specific developmental niche of extended immaturity which can scaffold the development of complex human skills, such as language and perspective-taking. Pretend play serves as one such scaffold in this developmental niche. It serves as a training ground for the acquisition of complex linguistic skills and skills for linguistically negotiating perspectives, but increased linguistic skills also afford more complex forms of pretend play, creating a feedback loop. As such, they also help children to practice and internalize complex cultural practices and roles. In this paper, it is argued that the cultural practice of pretend play within the human developmental niche represents an important factor that scaffolded the evolution of language.

1. Introduction

Pretend play seems to be a uniquely human behaviour that is culturally universal and displays a predictable developmental sequence (Lillard 2017). Pretend activities also make up a significant amount of children's daily interactions (Haight & Miller 1993; Hofferth & Sandberg 2001). This has prompted many researchers to propose that pretend play has a crucial role in children's development (e.g., Bergen 2002). Indeed, pretend play has been found to be closely connected and tightly integrated with other uniquely human cognitive and interactional abilities. For example, pretend play has been positively related to Theory of Mind, executive functions, and advanced sociocognitive capacities, especially in the form of pretend social role play (e.g. Carlson & White 2013). Pretend play is also strongly associated with language and language acquisition (Quinn et al. 2018). Given this relationship, some researchers also have assigned an important role to (pretend) play as a factor influencing the evolution of language (e.g. Lillard 2017; Langley et al. 2019). In Parker's (2002: xv) words, “[g]iven the developmental and evolutionary proximity between pretense and early language, perhaps it is inevitable that interest in the developmental and evolutionary emergence of language lurks behind much of the work on pretense.“ This paper will discuss the evolutionary foundations and functions of pretend play, especially its possible relationship to the evolution of language. I will first discuss the evolutionary foundations and functions of (pretend) play, before turning to the function of pretend play in human development. One of the key functions of pretend play is that it serves as a scaffold for development of social and cognitive abilities, including language. This is why I will look at the relationship of pretend play in language in more detail, before arriving at a theoretical proposal for role of pretend play in language evolution. Specifically, I will argue that pretend play served as developmental niche which scaffolded the emergence of complex forms of language, namely the development of complex constructions for negotiating and sharing perspectives.

2. The Evolutionary Foundations and Functions of (Pretend) Play

From an evolutionary standpoint, it is an important question which cognitive capacities children engaging in pretend share with other animals and what the evolutionary foundations of pretend play are. First of all, play can be found in all primates. Pellegrini et al. (2007: 272) state that for primates, play can be seen as a prolonged phase of free exploration, providing “a low cost way to develop alternate responses to new and challenging environments” (cf. Kavanaugh 2011: 296).

Regarding object play, it seems that human children and wild young chimpanzees engage in it to a similar degree (15% for human children vs 10% for young chimpanzees; Ramsey & McGrew 2005; Callaghan & Corbit 2015: 271). Social play also occurs in non-primate species (Palagi 2011: 71). Pellis and Pellis (2011), for example, argue that in rats, early social play has important positive effects on the development of the coordination of social interactions and emotional regulation. They hypothesise that social play in young human children might serve a similar function (Callaghan & Corbit 2015: 271). Lillard (2015: 442f.), also poses that there might be an evolutionary connection between pretence and the play fighting observed across many animal species. She argues that both behaviours create a frame in which actions possess meanings that not identical with their meanings and consequences in the real world. Therefore, some kind of boundary that separates real and pretend must be maintained by animals when they engage in play fighting. There might therefore be some evolutionary foundations connecting human and non-human play and also pretend play in humans. However, the evolutionary functions of pretend play are less clear. Lillard (2015. 459) holds that we still do not know why children engage in pretend play or why they engage in different forms of play at all. However, the ontogeny of pretend play indicates that it is an evolved behaviour. Namely, there is a predictable developmental sequence to it, and it appears universally in all cultures, albeit in different expressions and with different frequencies (Lillard 2017).

Lillard (2017) hypothesises that pretend play might have been an exapted by-product of animal play fighting. Play fighting probably evolved in animals as it represented a way to practice and rehearse fighting skills. “Play fighting involves signalling that one is only playing, and these signals and the accompanying play acts share the structure of other symbolic acts.” (Lillard 2015: 459). Pretend play in human children and play fighting in animals can be seen as analogous as they share a number of isomorphic properties: “both involve an as-if world, reading signals that indicate this as-if status, and understanding that behaviors and objects in the as-if world stand for or are symbolic of behaviors and objects in the real world” (Lillard 2017). Both play fighting and pretend play therefore involve metacommunication and symbolic relationships.

Human caregivers use specific cues to signal pretend, both behaviourally and later linguistically (Nakamichi 2015). Many other animals, especially other mammals, also have ritualised ways to signal to conspecifics that their behaviour is pretend – i.e. that it is different from real fighting. For example, play fighting rats use ultrasonic, high-pitched vocalisations that signal that they are play fighting and also nuzzle a different area of their conspecific than if they would really bite them (Pellis & Pellis 2011; 2017). Dogs use ‘play bows’ as a signal that they are play fighting (Bekoff 1995) and primates such as chimpanzees, gorillas, baboons, and macaques use a so-called ‘play face’ (Liebal et al. 2014: 137f.) to indicate that they are not engaging in actual aggression (Lillard 2017). However, the claim that the structure of these acts can be seen as symbolic is controversial, as is the claim that non-human animals perform symbolic play at all (Callaghan & Corbit 2015: 270; Mitchell 2002).

Play fighting is also positively related to the development of executive functions in a range of animals, especially self-regulation (Pelis & Pelis 2017). It is also positively related to the development of social skills (Gray 2019). As mentioned above, there is also evidence that the development of these capacities is also supported by pretend play (Pellis & Pellis 2017; Lillard 2017).

However, at the moment the evidence on proposed evolutionary benefits and functions of play is still far from conclusive (cf. Sharpe 2019). Regardless, as the above discussion has shown, there is support for the claim that play behaviour represents a shared evolutionary platform for the evolution of human symbolic behaviour, including language and pretence. A number of researchers have indeed highlighted the role of pretend play in the evolution of language. In accordance with Piaget (e.g. 1962), who argued for a common origin of language and symbolic play, these researchers argue that both language and pretend play require similar representational and sociocognitive capacities. Therefore, pretend play might have played an important co-evolutionary role in the evolution of language (e.g., Donald 1991; Knight 1998, Ginsburg & Jablonka 2014). To investigate this issue further, in the next section we will turn to proposed functions of pretend play in human ontogeny.

3. The Functions of (Pretend) Play in Development

The functions of (pretend) play in human ontogeny is a complex and controversially discussed issue (e.g. Lillard et al. 2013; Harris & Jalloul 2013). It has been argued to be important in in cognitive development in a number of domains. For instance, pretend play has been shown to be positively related with developments of children's executive and cognitive functions such as inhibition, working memory, cognitive flexibility, planning, problem-solving, metacognition, self-regulation, counterfactual reasoning and decision making (e.g., Bergen 2002; Whitebread & O'Sullivan 2012; Carlson & White 2013). The enactment of imagination requires can open new search spaces for solutions to problems and for testing hypotheses (Langley et al. 2019). Pretend play can therefore be said to facilitate children's learning by allowing them to explore reactions to various situations and to practice and internalise behaviours (Gaskins 2013).

What follows from this is that, crucially, pretend play has important social functions, as it enables them to learn about stereotypical situations and roles that are socioculturally salient (Gaskins 2013). Research on cultural variation in pretend play has shown that pretend play universally serves the function to practice and internalize culturally salient frames, scripts, schemas and routines with the aid of linguistic interaction (Gaskins 2013). In fact, in hunter-gatherer groups play time is correlated with the likelihood the child will perform a particular role in adulthood, and the difficulty of the skill they are pretending to perform.

Given its fundamentally social nature, pretend play can also be linked to sociocognitive development. For example, it has been argued that one of the functions of pretend play is that it aids perspective-taking abilities and the internalisation of different perspectives (Vygotsky 1978; Robson 2012).

Capacities for theory of mind and social understanding have also been related to pretend play (Youngblade & Dunn 1995). Shared pretense also requires the coordination and negotiation of roles and the co-creation of a shared ‘we-perspective’ (Tuomela 2002). For this reason, Rakoczy (2006) has argued that pretend play can be seen as a crucial cradle of the development of shared intentionality, that is, the capacity to engage in shared cooperative activities with others with shared intentions (Tomasello 2008). Both perspective-taking and shared intentionality are also strongly implicated in language acquisition (e.g. Tomasello 2003, 2008) and it therefore not surprising that pretend play, language acquisition, and semiotic development are also closely correlated (Zlatev & McCune 2014; Quinn et al. 2018).

This is also where the relationship between pretense and the evolution of language comes into play. Specifically, by looking at the ontogenetic role of pretend play in language acquisition we can make inferences about their relationship in the evolution of language.

4. Pretend Play and Language

There is a number of ways that pretend play can be said to boost and aid in language development. In general, it can be said that pretend play serves a scaffolding function for the development of complex linguistic constructions for the negotiation of perspectives. This is mainly due to two reasons (see also Langley et al. 2019).

First, as seen in the previous section, pretend play scaffolds the development of a number of cognitive abilities and mechanisms which are also operative in language acquisition, such as social understanding, shared intentionality, abstraction and schematization, (e.g. Tomasello 2003). Symbolic ability plays a particular important role in this context. As outlined in Section 2, both pretend play and language depend on the capacity for symbolic understanding (Lillard 2017; Zlatev & McCune 2014). That is, they both depend on the capacity to see one entity as symbolically standing in for and evoking another entity. In the case of pretend play, this could be the symbolic relationship between a banana and a telephone, where the banana is used ‘as-if’ it were a telephone (Lillard 2017). In language, it relates to the basic symbolic relationship between a linguistic form, like dog, and the conceptual entity it evokes or expresses (Langacker 1987). Recent evidence lends support to the proposal that both language and pretend play depend on the development of a symbolic capacity, which also can be seen as the foundation of other capacities related to pretend play and language, such as theory of mind (Lillard & Kavanaugh 2014).

Second, this is the case because the complexity of play and its cognitive and interactive elements foster the development of strategies negotiating complex pretend play scenarios, including the coordination of pretend actions and the assignment of pretend roles. For example, Cook-Gumperz & Kyratzis (2001) have shown that pretend play situations can be seen as a training ground and crucial scaffolding for the acquisition of progressive and simple present constructions. Overall, "Symbolic play provides a rich context for the exchange and negotiation of meaning, and thus may contribute to the development of important skills underlying communicative development" (Quinn & Kidd 2019: 33). Similarly, Trawick-Smith (1998: 433) notes that many researchers see pretend play as an important context for children to acquire linguistic competence and social skills because it requires them to "regularly negotiate shared symbolic meanings and coordinate ideas and intentions within make believe." This is echoed by Bruner (1983: 65), who argues that "the most complicated grammatical and pragmatic forms of language appear first in play activity." Relatedly, Langley et al. (2019) argue that play situations, due to their interactional complexity can be said to provide children with affordances and contexts for practice, hypothesis-testing and inductions of complex grammatical constructions. There is indeed evidence that children's use of language is more complex in play than in non-play contexts (Weisberg et al. 2013: 43). Internal state language, too, such as "This is a bad dog, you know" (Howe et al. 2005), as well as the linguistic co-construction of shared meanings have been shown to be positively related with pretend play (Howe et al. 2005). Pretend play, then, can be seen as a training ground and scaffolding for language acquisition and the linguistic negotiation of perspectives.

Given the developmental relationship between pretend play and the acquisition of more complex forms of language it is likely that they co-evolved. We will turn to this topic in the next section.

5. Pretend Play and Language Evolution

The above considerations allow us to draw several lines of argumentation together in order to arrive at a theoretical explication of the relationship of pretend play and language evolution. Both pretend play and language can be said to have been integrated into a human developmental niche through processes of niche construction and biocultural enculturation. This niche was created through the evolution of an extended juvenile period which in turn shaped human evolution and the evolution of language (cf. Sinha 2009). This extended period of socialization, in combination with more complex social networks had a number of effects on human ontogeny (Hare 2017; Benítez-Burraco & Kempe 2018). It can therefore be seen as an important component part of the evolution of the human socio-cognitive niche (Whiten & Erdal 2012).

Most significantly, in this niche children had more time to develop complex skills and cognitive abilities (Bjorklund et al. 2009). As a consequence, behaviours such as pretend play could scaffold other more complex behaviours such as the acquisition of complex constructions and skills for participating in complex culture. In addition, linguistically mediated play activities also scaffolded and were scaffolded by the emergence of complex skills of perspective-taking and social understanding (e.g. Tomasello 2008; Carpendale & Lewis 2015). Negotiating perspectives is central for human interaction (e.g. Carpendale & Lewis 2015) and construal and perspective-taking are also central to the structure of human language (e.g. Langacker 1987). As pretend play is also characterized by high degrees of negotiating perspectives, I argue that pretend play as a developmental niche and cultural practice scaffolded the emergence of complex constructions for negotiating, sharing, and maintaining perspectives.

One crucial factor in the evolution of language therefore was the emergence of a developmental niche in which pretend play scaffolded the acquisition and development of complex linguistic constructions in order to negotiate and share perspectives and to internalize culturally salient roles and knowledge. The development of more complex constructions afforded more complex forms of pretend play, which in turn led to the internalization of more complex cultural practices and mastery of cultural artefacts. In other words, humans constructed a developmental niche in which pretend play and language scaffolded each other in spiralling dynamic feedback loops within a multidimensional developmental web (e.g., Caporael et al. 2014; Mascolo and Fischer 2015).

One factor hypothesized to have led to the emergence of human forms of play is that of self-domestication (Langley et al. 2019, see also Benítez-Burraco & Kempe 2018). Domestication increases play behaviour in animals (Himmler et al. 2013) and leads to an extended juvenile period characterized by immaturity. It is likely that human self-domestication process also led to the entrenchment of longer and different types of play behaviour, in turn affording scaffolding processes enabling the development of complex linguistic skills and interactive behaviours (cf. Benítez-Burraco & Kempe 2018). This is consistent with arguments that the cultural evolution of linguistic structure is based on processes of self-domestication (Thomas & Kirby 2018).

6. Conclusion

In this paper I have argued that pretend play had an important role in the evolution of language and cognition. Play behaviour is present in a wide range of animals and has deep evolutionary roots. Both in human and non-human animals play can be said to scaffold the development of complex cognitive and social skills. The beginnings of symbolic abilities can also be seen in play behaviour. In humans, pretend play has an even more dramatic effect, and scaffolds the development of a wide range of cognitive and cultural skills. Importantly, it can be said to act as a scaffold for the development of perspective-taking skills and complex linguistic skills, which in turn feed back into the development of more complex forms of pretend play. For the evolution of language, this means that pretend play likely served as a scaffold enabling the development of complex constructions for negotiating perspectives, and for the linguistically mediated acquisition of cultural knowledge. This scaffolding process represents an emergent product of processes of self-domestication and the construction of a particular developmental niche in which this dynamic, interactive feedback loop could take place.

References

Bekoff, M. (1995). Play signals as punctuation: The structure of social play in canids. *Behaviour, 132*(5-6), 419-429.

Benítez-Burraco, A., & Kempe, V. (2018). The emergence of modern languages: has human self-domestication optimized language transmission? *Frontiers in Psychology, 9*, 551.

Bergen, D. (2002). The role of pretend play in children's cognitive development. *Early Childhood Research & Practice, 4*(1), n1.

Bjorklund, D. F., Periss, V., & Causey, K. (2009). The benefits of youth. *European Journal of Developmental Psychology, 6*(1), 120-137.

Bruner, J. (1983). *Child's talk: Learning to Use Language*. Oxford: Oxford University Press.

Callaghan, T., & Corbit, J. (2015). The Development of Symbolic Representation. In L. S. Liben & U. Müller (Eds.), *Handbook of Child Psychology and Developmental Science. Volume 2: Cognitive Processes* (7th ed., pp. 250-295). Hoboken, NJ: Wiley.

Carlson, S. M., & White, R. E. (2013). Executive Function, Pretend Play, and Imagination. In M. Taylor (Ed.), *The Oxford Handbook of the Development of Imagination* (pp. 161-174). Oxford: Oxford University Press.

Carpendale, J. I. M., & Lewis, C. (2015). The Development of Social Understanding. In L. S. Liben & U. Müller (Eds.), *Handbook of Child Psychology and Developmental Science. Volume 2: Cognitive Processes* (7th ed., pp. 381-424). Hoboken, NJ: Wiley.

Carporael, L. R., Griesemer, J. R., & Wimsatt, W. C. (Eds.). (2014). *Developing Scaffolds in Evolution, Culture, and Cognition*. Cambridge, MA: MIT Press.

Cook-Gumperz, J., & Kyratzis, A. (2001). Pretend play: Trial ground for the simple present. In M. Pütz, S. Niemeier, & R. Dirven (Eds.), *Applied Cognitive Linguistics I: Theory and Language Acquisition* (pp. 41-62). Berlin: De Gruyter.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Cambridge, MA: Harvard University Press.

Gaskins, S. (2013). Pretend Play as Culturally Constructed Activity. In M. Taylor (Ed.), *The Oxford Handbook of the Development of Imagination* (pp. 224-247). Oxford: Oxford University Press.

Ginsburg, S., & Jablonka, E. (2014). Memory, imagination, and the evolution of modern language. In D. Dor, C. Knight, & J. Lewis (Eds.), *The Social Origins of Language* (pp. 317-324). Oxford: Oxford University Press.

Gray, P. (2019). Evolutionary Functions of Play: Practice, Resilience, Innovation, and Cooperation. In P. K. Smith (Ed.), *The Cambridge handbook of play: Developmental and disciplinary perspectives* (pp. 84-102). Cambridge: Cambridge University Press.

Haight, W. L., & Miller, P. J. (1993). *Pretending at home: Early development in a sociocultural context*. Albany, NY: SUNY Press.

Hare, B. (2017). Survival of the friendliest: Homo sapiens evolved via selection for prosociality. *Annual Review of Psychology, 68*, 155-186.

Harris, P. L., & Jalloul, M. (2013). Running on Empty? Observing Causal Relationships of Play and Development. *American Journal of Play, 6*(1), 29-38.

Himmler, B. T., Stryjek, R., Modlinska, K., Derksen, S. M., Pisula, W., & Pellis, S. M. (2013). How domestication modulates play behavior: A comparative analysis between wild rats and a laboratory strain of Rattus norvegicus. *Journal of Comparative Psychology, 127*(4), 453.

Howe, N., Petrakos, H., Rinaldi, C. M., & LeFebvre, R. (2005). “This is a bad dog, you know…”: Constructing shared meanings during sibling pretend play. *Child Development, 76*(4), 783-794.

Knight, C. (1998). Ritual/speech coevolution: a solution to the problem of deception. In J. R. Hurford, M. Studert-Kennedy, & C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 68-91). Cambridge: Cambridge University Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Langley, M.C., A. Benítez-Burraco, & V. Kempe. (2019) Playing with language, creating complexity: Has play contributed to the evolution of complex language? *Evolutionary Anthropology*. https://doi.org/10.1002/evan.21810.

Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2014). *Primate Communication: A Multimodal Approach*. Cambridge: Cambridge University Press.

Lillard, A. S. (2015). The Development of Play. In L. S. Liben & U. Müller (Eds.), *Handbook of Child Psychology and Developmental Science. Volume 2: Cognitive Processes.* (7th ed., pp. 425-468). Hoboken, NJ: Wiley-Blackwell.

Lillard, A. S. (2017). Why Do the Children (Pretend) Play? *Trends in Cognitive Sciences, 21*(11), 826-834. doi:http://dx.doi.org/10.1016/j.tics.2017.08.001

Lillard, A. S., & Kavanaugh, R. D. (2014). The contribution of symbolic skills to the development of an explicit theory of mind. *Child Development, 85*(4), 1535-1551.

Lillard, A. S., Lerner, M. D., Hopkins, E. J., Dore, R. A., Smith, E. D., & Palmquist, C. M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin, 139*(1), 1.

Mascolo, M. F., & Fischer, K. W. (2015). Dynamic development of thinking, feeling, and acting. In W. F. Overton & P. C. M. Molenaar (Eds.), *Handbook of child psychology developmental science. Volume 1: Theory and Method* (7th ed., pp. 113-161). Hoboken, NJ: Wiley-Blackwell.

Mitchell, R. W. (Ed.) (2002). *Pretending and imagination in animals and children*. Cambridge: Cambridge University Press.

Nakamichi, N. (2015). Maternal behavior modifications during pretense and their long-term effects on toddlers' understanding of pretense. *Journal of Cognition and Development, 16*(4), 541-558.

Palagi, E. (2011). Playing at Every Age: Modalities and Potential Functions in Non-Human Primates. In A. D. Pellegrini (Ed.), *The Oxford Handbook of the Development of Play* (pp. 70-82). Oxford: Oxford University Press.

Parker, S. T. (2002). Foreword. In R. W. Mitchell (Ed.), *Pretending and Imagination in Animals and Children* (pp. xiv-xvi). Cambridge: Cambridge University Press.

Pellegrini, A. D., Dupuis, D., & Smith, P. K. (2007). Play in evolution and development. *Developmental review, 27*(2), 261-276.

Pellis, S. M., & Pellis, V. C. (2011). Rough-and-tumble play: training and using the social brain. In A. D. Pellegrini (Ed.), *The Oxford Handbook of the Development of Play* (pp. 245-259). Oxford: Oxford University Press.

Pellis, S. M., & Pellis, V. C. (2017). What is play fighting and what is it good for? *Learning & Behavior, 45*(4), 355-366. doi:10.3758/s13420-017-0264-3

Piaget, J. (1962). *Play, dreams and imitation in childhood*. New York: Norton.

Quinn, S., Donnelly, S., & Kidd, E. (2018). The relationship between symbolic play and language acquisition: A meta-analytic review. *Developmental review, 49*, 121-135.

Quinn, S., & Kidd, E. (2019). Symbolic play promotes non-verbal communicative exchange in infant-caregiver dyads. *British Journal of Developmental Psychology, 37*(1), 33-50.

Rakoczy, H. (2006). Pretend play and the development of collective intentionality. *Cognitive Systems Research, 7*, 113-127.

Ramsey, J., & McGrew, W. C. (2005). Object Play in Great Apes: Studies in Nature and Captivity. In A. D. Pellegrini & P. K. Smith (Eds.), *The Nature of Play* (pp. 89-112). New York: The Guilford Press.

Robson, S. (2012). *Developing thinking and understanding in young children: An introduction for students* (2nd ed.). London: Routledge.

Sharpe, L. (2019). Fun, Fur and Future Fitness: The Evolution of Play in Mammals. In Peter K. Smith & J. L. Roopnarine (Eds.), *The Cambridge handbook of play: Developmental and disciplinary perspectives* (pp. 49-66). Cambridge: Cambridge University Press.

Sinha, C. (2009). Language as a biocultural niche and social institution. In V. Evans & S. Pourcel (Eds.), *New directions in cognitive linguistics* (pp. 289-310). Amsterdam & Philadelphia: John Benjamins.

Thomas, J., & Kirby, S. (2018). Self domestication and the evolution of language. *Biology & philosophy, 33*(1-2), 9.

Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge, MA: MIT Press.

Trawick-Smith, J. (1998). A qualitative analysis of metaplay in the preschool years. *Early Childhood Research Quarterly, 13*(3), 433-452.

Tuomela, R. (2007). *The Philosophy of Sociality: The Shared Point of View*. Oxford: Oxford University Press.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman Eds.). Cambridge, MA: Harvard University Press.

Weisberg, D. S., Zosh, J. M., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Talking It Up: Play, Language Development, and the Role of Adult Support. *American Journal of Play, 6*(1), 39-54.

Whitebread, D., & O'Sullivan, L. (2012). Preschool children’s social pretend play: Supporting the development of metacommunication, metacognition and self-regulation. *International Journal of Play, 1*(2), 197-213.

Whiten, A., & Erdal, D. (2012). The human socio-cognitive niche and its evolutionary origins. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1599), 2119-2129.

Youngblade, L. M., & Dunn, J. (1995). Individual differences in young children’s pretend play with mother and sibling: Links to relationships and understanding of other people's feelings and beliefs. *Child Development, 66*(5), 1472-1492.

Zlatev, J., & McCune, L. (2014). Toward and integrated model of semiotic development.

LEXICAL AND GESTURAL ALIGNMENT IN INTERACTION AND THE EMERGENCE OF NOVEL SHARED SYMBOLS

MARLOU RASENBERG[*1,2], MARK DINGEMANSE[1,2], and ASLI ÖZYÜREK[1,2,3]

[*]Corresponding Author: marlou.rasenberg@mpi.nl
[1]Centre for Language Studies, Radboud University, Nijmegen, The Netherlands
[2]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands

A challenge in the field of language evolution and emergence of new languages is to explain how shared symbols are created for concepts for which conventional labels do not exist. Converging on shared labels can be needed even in interactions between people who do share a conventional language but need to align their concepts for mutual understanding (e.g., when talking about new ideas, plans, innovations or designs). Here we investigate the role of multimodal (gestural and lexical) alignment – that is, using multimodal labels with similar semantic and form features during collaborative referring – in the creation of shared symbols for novel referents among dyads.

Previous research has shown that repetition of lexical items plays an important role in collaborative referring, be it through incremental grounding of shared labels (e.g., Brennan & Clark, 1996) or priming (e.g., Pickering & Garrod, 2004). Gestures too can be a powerful interactional resource, as they allow for iconic mappings and can convey meaning beyond what is expressed in speech (McNeill, 1992). Like words, gestures can be repeated (or “mimicked”), and such repetition has also been shown to aid collaborative referring (Holler & Wilkin, 2011). An open question is the relation between alignment in these two modalities in the context of creating shared labels for novel referents, which remains understudied (cf. Oben & Brône, 2016).

To investigate patterns of multimodal alignment in interaction and the creation of novel labels, we used a director-matcher task in which dyads communicate about unfamiliar 3D objects (i.e., ‘Fribbles’, Barry, Griffith, De

Rossi, & Hermans, 2014) in 6 consecutive rounds. For this *interaction task*, speech has been transcribed and co-speech gestures annotated. Before and after this interaction, participants individually labelled each object (in 1-3 words) for their partner in a *naming task*. For the purposes of this paper we analyzed data from 8 dyads, 8 target objects and the first two rounds of the interaction task.

To see if dyads converged in their labels for objects before and after the interaction in the *naming task*, we used pretrained word2vec word embeddings (based on the NLPL Dutch CoNLL17 corpus), and quantified the semantic similarity of these descriptions, yielding a simple measure of relative convergence in naming practices. There was a significant increase in the semantic similarity of descriptions, when comparing pre-interaction naming (M = 0.48, SD = 0.16) with post-interaction naming (M = 0.69, SD = 0.24); t(63) = -5.80, p < .001 (one-tailed). In contrast, dyads of participants who did not interact with each other did not show similar post-interaction convergence.

As a next step, the *interactional task* was analyzed to see if lexical and gestural alignment occurred and if they correlated, possibly giving rise to post-interaction convergence. Among the extensive use of iconic gestures (N=561 for 8 dyads), more than half (66.5%) was found to be ‘referentially aligned’; that is, these gestures depicted the same subparts of the target objects. Within this set of referentially aligned gesture pairs, the degree of form alignment was calculated by scoring overlap in handedness, position, handshape, orientation and movement. Full alignment of all five form features was uncommon in these dyads (only 4% of all cases), while partial alignment of one up to four features occurred frequently (90%). Subsequently, for those target objects where referential alignment of gestures occurred for a given dyad, a lexical alignment score for that target was computed by comparing the referential speech from both participants (yielding a cosine similarity score). For these cases, a moderate correlation between the number of referentially aligned gestures and lexical alignment was found ($r_s = .39$; p = .010).

These results show that when creating shared labels for novel objects, interactants align in their iconic gestures as well as their speech, and these two types of alignment are correlated with each other. However, while the presence of gestural alignment is evident at the level of semantics, full form alignment does not seem to be a necessary feature. Existence of partial alignment in gesture is in line with previous reports of grounding processes for speech used in interaction (Clark & Wilkes-Gibbs, 1986). These findings underscore the role of interaction and strategic multimodal alignment as important resources for the emergence of novel shared symbols.

Acknowledgements

This work was supported by the Dutch Research Council (NWO) Gravitation Grant 024.001.006 to the Language in Interaction Consortium and Vidi grant 016.vidi.185.205 to MD. The study came about in the context of the multidisciplinary research project *Communicative Alignment in Brain & Behavior* (CABB) within the Language in Interaction consortium. The authors are thankful to CABB members (in alphabetical order) Mark Blokpoel, Sara Bögels, Lotte Eijk, Mirjam Ernestus, Judith Holler, Branka Milivojevic, Iris van Rooij, Ivan Toni, and Herbert Schriefers, amongst others, for their contributions in various stages of the study. We are also grateful to Emma Berensen, Junfei Hu and Clarissa de Vries for their help with coding and analyses. The authors remain solely responsible for contents of the abstract.

References

Barry, T. J., Griffith, J. W., De Rossi, S., & Hermans, D. (2014). Meet the Fribbles: Novel stimuli for use within behavioural research. *Frontiers in Psychology, 5,* 103.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482–1493.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Holler, J., & Wilkin, K. (2011). Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue. *Journal of Nonverbal Behavior*, *35*(2), 133–153.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago, IL: University of Chicago Press.

Oben, B., & Brône, G. (2016). Explaining interactive alignment: A multimodal and multifactorial account. *Journal of Pragmatics*, *104*, 32–51.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190.

NETWORK STRUCTURE AND THE CULTURAL EVOLUTION OF LINGUISTIC STRUCTURE: A GROUP COMMUNICATION EXPERIMENT

LIMOR RAVIV*[1], ANTJE MEYER[1,2], and SHIRI LEV-ARI[3]

*Corresponding Author: limor.raviv@mail.huji.ac.il
[1]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[2]Radboud University, Nijmegen, The Netherlands
[3] Royal Holloway University of London, Egham, UK

Social network structure has been argued to shape the structure of languages, as well as affect the spread of innovations and the formation of conventions in the community. Specifically, theoretical and computational models of language change suggest that sparsely connected communities develop more regular and more systematic languages, while tightly knit communities can maintain high levels of linguistic complexity (e.g., Meir, Israel, Sandler, Padden & Aronoff, 2012; Trudgill, 2009; Wray and Grace, 2007; Lou-Magnuson and Onnis, 2018). This is because sparser communities are typically more diversified (e.g., Bahlmann, 2014) and tend to converge more slowly (e.g., Zubek et al., 2017). Importantly, sparser networks' greater variability and convergence difficulty can trigger a stronger need for generalizations and regularizations, leading to the creation of more systematic languages (Raviv, Meyer & Lev-Ari, 2019; Wray and Grace, 2007). However, this hypothesis has not yet been tested experimentally.

The goal of the current study is to fill in this gap in the literature, and experimentally test the effect of network structure using a group communication paradigm (Raviv et al., 2019). We examined the formation of new languages developed by different micro-societies, in which participants interacted face-to-face about novel scenes in alternating pairs. Groups varied only in their network structure (Fig 1), while keeping community size constant: All networks were comprised of eight participants, yet differed in their degree of connectivity (i.e., how many people each participant interacted with) and homogeneity (i.e., whether all participants are equally connected). Specifically, we contrasted three types of networks (seven groups per condition), which are typically used in computational models and echo early and contemporary human societies:

- **Fully connected network (Fig 1A):** a maximally dense and homogenous network (global clustering coefficient = 1) where all participants interact.

- **Small-world network (Fig 1B):** a sparser network with only half of the possible connections (global clustering coefficient = 0.17). "Strangers" are indirectly linked via short paths (max 2).
- **Scale-free network (Fig 1C):** similar in sparsity to the small-world networks, except that its distribution of connections follows a power law: most agents have few connections, but some have many (“hubs”, e.g., node A; global clustering coefficient = 0.42).

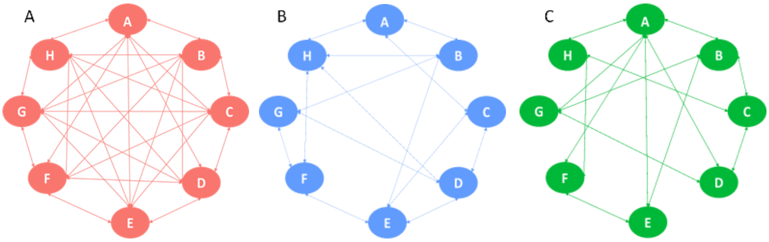


Figure 1. Network structure conditions. We tested group of eight participants who were connected to each other in different network setups: fully connected (A), small-world (B), and scale-free (C).

Across conditions, languages became more communicatively successful, more systematically structured (i.e., higher correlations between string distances and sematic distances; Fig. 2), more stable (i.e., fewer changes over time), and more shared (i.e., higher similarity between different participants’ variants) over time. There was no significant effect of network structure for any measure. We hypothesize that these null results can be traced back to the fact that, surprisingly, the networks did not differ in the amount of input variability. At the same time, small-world networks showed the greatest variation across all measures. This greater variability suggest that network structure can influence susceptibility to random linguistic changes (i.e., drift).

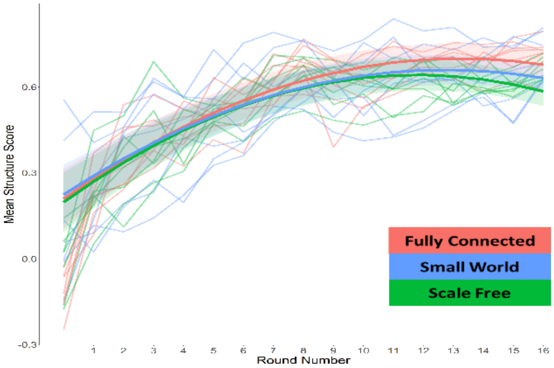


Figure 2. Changes in linguistic structure over time by network structure condition.

References

Bahlmann, M. D. (2014). Geographic Network Diversity: How Does it Affect Exploratory Innovation? *Industry and Innovation*, *21*(7–8), 633–654.

Lou-Magnuson, M., Onnis, L., 2018. Social Network Limits Language Complexity. Cognitive Science, 42, 2790–2817.

Meir, I., Israel, A., Sandler, W., Padden, C. A., & Aronoff, M. (2012). The influence of community on language structure: evidence from two young sign languages. Linguistic Variation, 12(2), 247-291.

Raviv, L., Meyer, A., Lev-Ari, S., 2019. Larger communities create more systematic languages. Proceedings of the Royal Society B: Biological Science, 286, 20191262.

Trudgill, P. (2009). Sociolinguistic typology and complexification. In G. Sampson, D. Gil, and P. Trudgill (eds.), Language Complexity as an Evolving Variable. Oxford: Oxford University Press.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. Lingua, 117(3), 543-578.

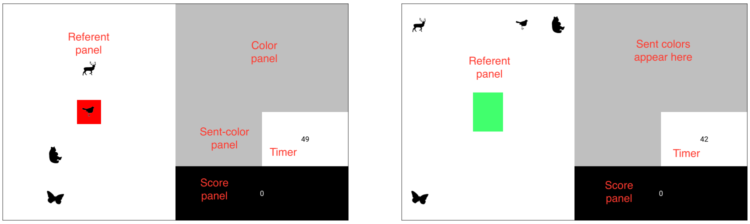
Zubek, J., Denkiewicz, M., Barański, J., Wróblewski, P., Rączaszek-Leonardi, J., & Plewczynski, D. (2017). Social adaptation in multi-agent model of linguistic categorization is affected by network information flow. *PLOS ONE*, *12*(8), e0182490.

THE EMERGENCE OF PHONOLOGY-LIKE ORGANIZATION IN A QUASI-PHONETIC SPACE

GARETH ROBERTS[*1] and ROBIN CLARK[1]

[*] Corresponding Author: gareth.roberts@ling.upenn.edu
[1] Department of Linguistics, University of Pennsylvania, Philadelphia, USA

It has long been observed that phonemes are not sampled at random from available phonetic space but are rather well dispersed (Boer, 2000; Lindblom, 1986; Lindblom & Maddieson, 1988). Vowel spaces are a good example. One hypothesized explanation is that well dispersed systems allow phonemes to be distinguished more easily (Liljencrants & Lindblom, 1972; Schwartz, Boë, Vallée, & Abry, 1997). However, phonemes located closer to the edges of the space are also easier for speakers to locate reliably. We conducted an experiment to investigate the role of interactive processes in the emergence of "phonological" organization in a novel communication medium, manipulating the extent to which perceptual demands aligned with production demands.



(a) Sender's screen

(b) Receiver's screen

Figure 1.: Screenshots from the game

60 undergraduates played a referential game in pairs. Players took turns to be *Sender* and *Receiver*, with the Sender communicating animal silhouettes. This involved moving a finger around on a trackpad, causing colors to appear in an onscreen *Color panel* (Fig. 1a). The Sender could select colors by holding their finger in place for 1s, causing the color to appear on the Receiver's screen (Fig. 1b) for 2s. The Sender could send as many colors as they wanted within 20s, and the Receiver had to select the right animal. Each color was composed of red, green, and blue components, the contribution of each ranging from 0 to 1. Fig. 2 shows

examples of how finger-position corresponded to colors (though neither player ever saw the whole space in this way; the relationship between specific xy dimensions and color-components was counterbalanced. Initially the pair communicated about four referents; if they successfully communicated each referent at least 75% of the time, four more were added, up to a total of 12. We manipulated how well production pressures aligned with perceptual pressures: In the *Outer-edge* condition, brighter easier-to-distinguish colors were located around the edge of the trackpad and were thus easier for the Sender to locate reliably. In the *Inner-edge*, the brightest colors were around a harder to locate "inner edge".

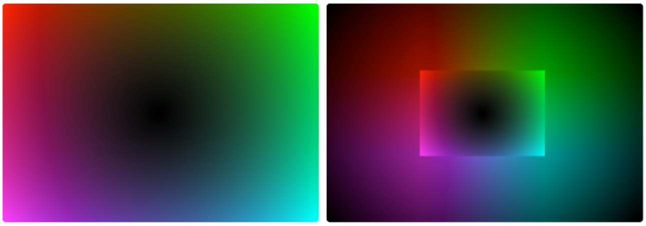


Figure 2.: Colorspaces for Outer-edge and Inner-edge conditions

We identified repeated "phonemes" across signals using Pillai scores (for a discussion of their use to identify merged and unmerged vowels in sociolinguistics, see Nycz & Hall-Lew, 2013). We then measured dispersion in terms of mean pairwise distance and distance from center, as well as mode brightness (the brightness of the brightest component in the RGB space). We calculated a success index as $(\sum_1^{n_r} s)/12n_r$, where n_r is the number of rounds and the numerator is a cumulative count of s, the number of successfully established "words" in a given round. Participants in the Inner-edge condition found the game harder and levels of dispersion in the "articulatory" space were significantly lower, suggesting that perception was guiding production; interestingly, however, mode brightness was also lower in the perceptual space, suggesting that participants were not simply maximizing perceptibility, but were having to find a compromise between perception and production demands. Dispersion in the production space was at greater than chance levels in the Outer-edge condition, but not the Inner-edge condition. Success was significantly correlated with dispersion across conditions, but not within conditions. This suggests success was driven not by dispersion per se, but by the alignment between the demands acting on the producer and the perceiver. This suggests we should expect the location of phonemes to be driven not by dispersion alone, but by the extent to which production and perception are mutually reinforcing, consistent with theoretical models in which the topology of the signaling space plays an important role (Stevens & Keyser, 2010; Carré, Divenyi, & Mrayati, 2017; Schwartz, Abry, Boë, Ménard, & Vallée, 2005)

References

Boer, B. de. (2000). Self-organization in vowel systems. *Journal of Phonetics*, *28*(4), 441–465.

Carré, R., Divenyi, P., & Mrayati, M. (2017). *Speech: A dynamic process.* Berlin/Boston: de Gruyter.

Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, *48*(4), 839–62.

Lindblom, B. (1986). Phonetic universals in vowel systems. *Experimental phonology*, 13–44.

Lindblom, B., & Maddieson, I. (1988). Phonetic universals in consonant systems. *Language, Speech, and Mind*, 62–78.

Nycz, J., & Hall-Lew, L. (2013). Best practices in measuring vowel merger. In *Proceedings of meetings on acoustics* (Vol. 20).

Schwartz, J.-L., Abry, C., Boë, L.-J., Ménard, L., & Vallée, N. (2005). Asymmetries in vowel perception, in the context of the Dispersion–Focalisation Theory. *Speech Communication*, *45*(4), 425–434.

Schwartz, J.-L., Boë, L.-J., Vallée, N., & Abry, C. (1997). The dispersion-focalization theory of vowel systems. *Journal of phonetics*, *25*(3), 255–286.

Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, *38*(1), 10–19.

DOES GREATER USE OF LANGUAGE PROMOTE GREATER CONCEPTUAL ALIGNMENT?

HETTIE ROEBUCK [*1], GARY LUPYAN[1]

[*]Corresponding Author: hroebuck@wisc.edu
[1]Department of Psychology, UW-Madison, Madison, USA

People show a surprisingly high degree of conceptual alignment even in the face of sometimes large differences in perceptual experience (Dieciuc & Folstein, 2019). For example, when asked to generate features of common objects, sighted and congenitally blind people show near perfect agreement ($r= 0.96$) (Lenci, Baroni, Cazzolli, & Marotta, 2013; Bedny, Koster-Hale, Elli, Yazzolino, & Saxe, 2019; Kim, Elli, & Bedny, 2019). People also show quite high agreement in more open-ended tasks people, e.g., asked to draw a piece of furniture (one of the current studies), 47% of people drew a chair.

Here, we ask whether conceptual alignment is increased by the use of language. Why would language promote alignment? At its most basic, language provides us with a set of category labels which group together distinct perceptual experiences. This applies across concrete concepts: an ant and a butterfly might both be labeled as “insects”, to more abstract concepts: a painting on a wall and an apple on a table are labeled as instances of “on”-ness (Bedny et al., 2019; Kim et al., 2019). Different experiences, when labeled using common categorical terms, become represented as more similar and because labels highlight category-diagnostic dimensions, may become represented as more typical (Lupyan, 2012).

Examining links between language and conceptual alignment is relevant for understanding the evolution of language because it helps us understand the *functions* of language. To the extent that greater conceptual alignment facilitates both communication and cooperation (Markman & Makin, 1998; Silvey, Kirby, & Smith, 2019), finding that language itself increases conceptual alignment hints that such an increase may have acted as a selection pressure on the emergence of language.

Here, we test the link between language and conceptual alignment in a novel way: by taking advantage of natural variability in people’s self-reported internal verbalization. While most people describe frequently experiencing inner speech, and consciously relying on language while planning, problem-solving, and recalling memories, others do so to a much lesser degree, or not at all (e.g., Alderson-Day & Fernyhough, 2015; Hurlburt, Heavey, & Kelsey, 2013).

We used a questionnaire to quantify differences in people's internal verbalization using Likert-type questions such as 'If I am walking somewhere by myself, I often have a silent conversation with myself.' Our questionnaire shows high reliability and predicts a number of objective difference in spontaneous uses of language (Roebuck & Lupyan, 2019). We then correlated this measure of *internal verbalization* with two measures of conceptual alignment. In Study 1, we asked people (64 adult English speakers) to rank photographs from common categories on their relative typicality, and then assessed whether the similarity of their rankings (a proxy for conceptual alignment), is predicted by their internal verbalization score. In Study 2, we presented people (n=200) with four drawing prompts (*piece of furniture*, *kitchen appliance*, *dessert*, *reptile*) and then recruited new people to rate each drawing on its typicality of the prompt category. We then used a convolutional neural network to measure visual similarity between each pair of images drawn from the same cue (e.g., all reptile drawings).

Study 1 found that people with higher internal verbalization rated the typicality of photographs (e.g., types of cars, dogs, etc.) in more similar ways to one another. Study 2 showed that people with higher internal verbalization produced drawings that were more typical exemplars of the cued category (Fig. 1). Although internal verbalization did not predict *which* basic-level types people drew when prompted with a superordinate cue (i.e., lizard vs. snake vs. turtle), higher internal verbalization was associated with drawing more typical lizards, snakes etc. (even when controlling for how well-drawn the image was *t*=2.0). When we compared pairs of drawings (e.g., every cake with every other cake), we found that more *similar* internal verbalization between drawers was associated with more similar ratings of typicality and more visual similarity between drawings (*t*=3.2). Taken together, the results support the idea that greater use of internal language promotes conceptual alignment, even for very familiar and concrete categories.

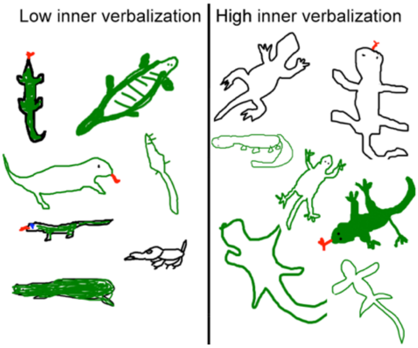


Fig 1 Example of lizard drawings from the *draw a reptile* prompt from people with low vs. high internal verbalization scores. Greater internal verbalization was associated with greater measured category typicality.

Our finding of lower conceptual alignment among people who report relying less on language offers a provocative hint at how much lower our alignment might be if we could not rely on language at all. Our ongoing work is testing the *causal* role of internal language in conceptual alignment by manipulating the availability of task relevant language.

Acknowledgements

This research was supported by NSF-PAC Grant 1734260 to GL.

References

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin*, *141*(5), 931–965. https://doi.org/10.1037/bul0000021

Bedny, M., Koster-Hale, J., Elli, G., Yazzolino, L., & Saxe, R. (2019). There's more to "sparkle" than meets the eye: Knowledge of vision and light verbs among congenitally blind and sighted individuals. *Cognition*, *189*, 105–115. https://doi.org/10.1016/j.cognition.2019.03.017

Dieciuc, M. A., & Folstein, J. R. (2019). Typicality: Stable structures and flexible functions. *Psychonomic Bulletin & Review*, *26*(2), 491–505. https://doi.org/10.3758/s13423-018-1546-2

Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, *22*(4), 1477–1494. https://doi.org/10.1016/j.concog.2013.10.003

Kim, J. S., Elli, G. V., & Bedny, M. (2019). Knowledge of animal appearance among sighted and blind adults. *Proceedings of the National Academy of Sciences*, *116*(23), 11213–11222. https://doi.org/10.1073/pnas.1900952116

Lenci, A., Baroni, M., Cazzolli, G., & Marotta, G. (2013). BLIND: A set of semantic feature norms from the congenitally blind. *Behavior Research Methods*, *45*(4), 1218–1233. https://doi.org/10.3758/s13428-013-0323-4

Lupyan, G. (2012). What Do Words Do? Toward a Theory of Language-Augmented Thought. In *Psychology of Learning and Motivation* (Vol. 57, pp. 255–297). https://doi.org/10.1016/B978-0-12-394293-7.00007-8

Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. *Journal of Experimental Psychology: General*, *127*(4), 331–354. https://doi.org/10.1037/0096-3445.127.4.331

Roebuck, H., & Lupyan, G. (2019, May 9). The Internal Representations Questionnaire: Measuring modes of thinking. https://doi.org/10.31234/osf.io/euhcn

Silvey, C., Kirby, S., & Smith, K. (2019). Communication increases category structure and alignment only when combined with cultural transmission. *Journal of Memory and Language*, *109*, 104051. https://doi.org/10.1016/j.jml.2019.104051

THE EVOLUTION OF VOCATIVE MORPHOLOGY AS SHAPED BY THE TUNE

TIMO B. ROETTGER[1] and MARTÓN SÓSKUTHY[2]
*Corresponding Author: timo.b.roettger@gmail.com
[1]Department of Cognitive Science, University of Osnabrück, Germany
[2]Department of Linguistics, University of British Columbia, Canada

Human communication systems can express multiple types of meaning simultaneously such as propositional meaning and discourse-pragmatic meaning. To do so language users often exploit different aspects of the speech signal. While propositional content tends to be signaled primarily through *segmental* means (i.e. different combinations of consonants and vowels), its broader interpretation within the discourse context is often expressed *prosodically* by the modulation of pitch across the utterance (henceforth referred to as the *tune*, see Ladd 2008).[1]

Segmental and prosodic aspects of the signal are not independent of each other. For successful articulation and perceptual retrieval of the tune, the segmental carriers need to have a rich harmonic structure and high periodic energy (e.g. Barnes et al. 2014). The optimal carriers of the tune are therefore vowels. When there is a shortage of vowels, communicatively relevant tunes can be phonetically impoverished (e.g. Odé 2005), endangering the recovery of the intended meaning. These functional conflicts can lead to biases towards speech variants that optimize the transmission of the tune including the insertion of non-lexical vowels (e.g. Roettger 2017, Roettger & Grice 2019). The results of such systematic biases might then be diachronically reanalyzed as a grammatical marker.

The present paper will argue that vocative morphology is an example of such a grammaticalization process. Vocative constructions – used to call interlocutors or to attract /maintain the addressee's attention (e.g. Daniel & Spencer 2009) – are often characterized by specific tunes (e.g. Ladd 1978). Many languages also mark vocatives morphologically. Since the tunes associated with vocatives require 'tune-friendly' segmental carriers, one potential pathway for the diachronic emergence of such morphological markers is the grammaticalization of tune-driven intrusive vowels (Roettger & Grice 2019). Such markers would consist of a single vowel, and therefore we predict that vocatives should contain consonants less frequently than other grammatical markers.

We performed a large-scale literature search and assembled a cross-linguistic database of 101 languages (46 language families) that have been described with

---

[1] Note that neither is propositional meaning restricted to segmental contrasts, nor is discourse-pragmatic meaning restricted to prosody.

grammatical vocatives. We extracted segmental information about vocative markers and compared them to structural case markers (SCM) that overtly mark either the agent or patient. The results are shown in Fig. 1(B). The estimated posterior probability of consonants in SCMs is 0.85 (95% Credible Interval [0.62,0.97]), while in vocatives it is only 0.40 (95% CI [0.17,0.66]). The estimated difference is –0.45 (95% CI [–0.69,–0.18]). These patterns hold even when only considering suffixes (see Fig. 1C). Our results indicate that vocative markers are substantially more tune-friendly than other comparable grammatical markers, suggesting that many of them may have emerged from tune-driven adjustments to the segmental material.

These findings suggest that accounts of the evolution of linguistic systems must consider the expression of different communicative functions (Foulkes et al. 2018). Here we have focused on how the acoustic properties of prosodic patterns (used to convey pragmatic meaning) interact with segmental features, and how this interaction potentially shapes the emergence of morphological vocatives.

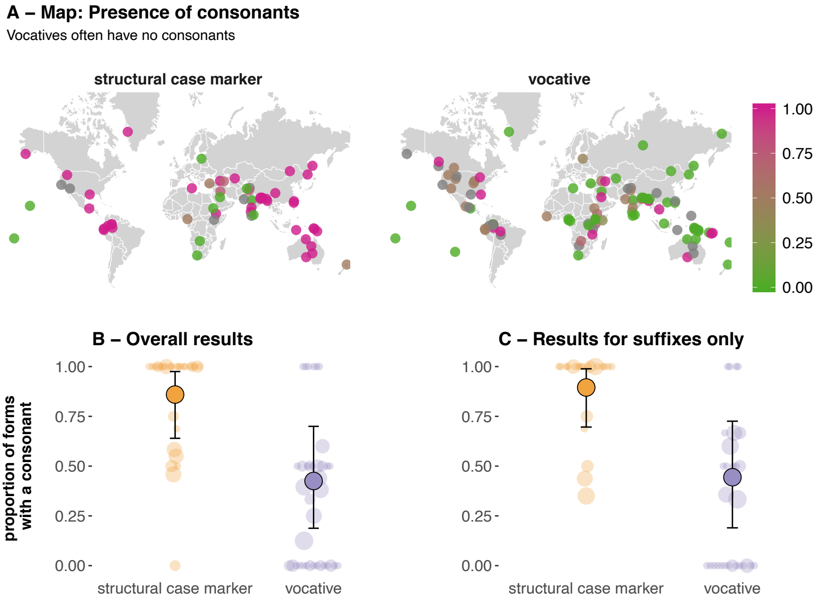


Figure 1. A - World map illustrating the proportion of morphological markers that have at least one consonant across the languages within our corpus; B - Model estimates (posterior means and 95% CIs) for all markers; C - Model estimates for only suffixes. Semi-transparent points are averages for language families scaled for the number of languages in each family.

References

Barnes, J., Brugos, A., Veilleux, N., & Shattuck-Hufnagel, S. (2014). Segmental influences on the perception of pitch accent scaling in English. *Proceedings of Speech Prosody*, 7, 1125–1129.

Daniel, M., & Spencer, A. (2009). The vocative–an outlier case. In A. Spencer & A. L. Malchukov (Eds.), *The Oxford handbook of case*. Oxford: Oxford University Press.

Foulkes, P., Docherty, G., Hufnagel, S. S., & Hughes, V. (2018). Three steps forward for predictability. Consideration of methodological robustness, indexical and prosodic factors, and replication in the laboratory. *Linguistics Vanguard*, 4(s2)

Ladd, D. R. (1978). Stylized intonation. Language, 54(3), 517–540.

Ladd, D. R. (2008). *Intonational Phonology* (2nd edition [1996]). Cambridge: Cambridge University Press.

Odé, C. (2005). Neutralization or truncation? The perception of two Russian pitch accents on utterance-final syllables. *Speech Communication*, 47(1–2), 71–79.

Roettger, T. B., & Grice, M. (2019). The tune drives the text - Competing information channels of speech shape phonological systems. *Language Dynamics and Change* 9(2), 265-298.

Roettger, T. B. (2017). *Tonal placement in Tashlhiyt: How an intonation system accommodates to adverse phonological environments*. Berlin: Language Science Press.

VISUAL RECURSION DEVELOPS IN ABSENCE OF LINGUISTIC RECURSION. A CASE REPORT

JOANA ROSSELLÓ*[1], ALEXANDRE CELMA-MIRALLES[2,3], MAURICIO DIAS MARTINS[4,5]

*Corresponding Author: joana.rossello@ub.edu
[1]Dpt. of Catalan Philology and General Linguistics, University of Barcelona, Spain
[2]Center for Brain and Cognition, Pompeu Fabra University, Barcelona, Spain
[3]Center for Music in the Brain, Aarhus University, Aarhus, Denmark
[4]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[5]Institut Jean Nicod, Département d'Etudes Cognitives, ENS, EHESS, CNRS, PSL Research University

Introduction

Recursion is a cognitive capacity, potentially unique to humans, which allows the generation of hierarchies with multiple levels of embedding. This capacity is thought to primarily underlie syntactic structures in language (Berwick & Chomsky, 2016), but is also available in other domains such as vision (Martins, 2012). An open question is whether a purportedly primary linguistic capacity is used in other domains, or whether visual recursion can develop in the absence of language. If the latter, is recursion a domain-general or multi-domain-specific?

To address this question, a Visual Recursion Task (VRT) has been used. In this task, participants are shown the first 3 steps of a recursive rule generating 3 visual fractals, and then asked to choose the correct 4th step from 2 alternatives. While in VRT each step generates a new hierarchical level, a control task —the Embedded Iteration Task (EIT)—, has been devised in which each step adds elements within a nested hierarchical level, without generating new levels.

Results suggest that VRT is not distinctively associated with language: (i) both visual recursion and iteration develop around 9-10 y.o. and both correlate with grammatical skills (Martins et al., 2014b), (ii) visual recursion is not interfered by a concurrent verbal memory task (Martins et al., 2015) and (iii) fMRI studies do not show language brain areas involved for either VRT or EIT (Martins et al., 2014a). However, these experiments could be criticized since, a) the grammar task in (i) did not contain complex sentences, b) the verbal memory task —digit sequence— does not tap into language specific resources (Cook et al., 2018), and c) the fMRI task was performed by well-trained participants, who might solve it via automatic visual template matching. Besides, a recent study (Martins et al., 2019) shows that lesions in the pMTG, a language area, impair visual recursion and that VRT, but not EIT, correlates with the comprehension of sentences with two levels of clausal center-embedding.

Case report

Here we further test the hypothesis that linguistic and visual recursion are independent, in the framework of a case study. Its subject is Álex (AX), a 12-year-old autistic child (ADI-r and ADOS assessed) with an oral open-ended but almost exclusively nominal lexicon in three languages, which he has learned mainly by reading, with nouns as captions of bi-dimensional images. He has no verbs and no grammar. Nouns for depicted objects together with some adjectives for perceptually salient attributes are not enough for AX to deploy a functional language. Similarly to nonverbal autistic children, AX‘s communication is only imperative even in the rare occasions he communicates orally. In such cases his utterances are maximally two-word demands (e.g. *batido rosa* Sp.: *pinky shake* Engl.). Lack of linguistic recursion in production extends to comprehension — AX’s PPVT-III/verbal mental age outcome, 4.3 y.o., manifestly overestimates his comprehension. He has a non-verbal IQ of 79, tested through Leiter-3 with the following subtest scores: Figure-Ground (FG), 5; Form Completion (FC), 5; Classification and Analogies (CA), 8; and Sequential Order (SO), 6. Only CA reaches the ‘medium’ range while the rest of subtests yield a ‘medium-low’ score. Interestingly, CA assesses pattern analysis and prediction of “what goes next” while FG is a visual interference task where the target object is embedded in an increasingly complex background; FC, a task where the subject has to arrange parts in a whole, demands a capacity for synthesis in visual organization and finally, SO, tests the ability to analyze sequential order (Roid & Koch, 2017).

We ran the EIT→VRT tests on AX, in this order, and his accuracy in VRT was similar to 9-10 y.o. age group (77%) (Martins, 2014b). Interestingly he was a slow responder in the first half of EIT (57% of trials timeouts) but performed well in the second half (77%).

Discussion

AX has no impairment in visual recursion despite lacking linguistic recursion, which suggests that linguistic recursion has no ontogenetic primacy. AX’s performance seems nevertheless contrary to the hypothesis of multiple-domain specificity because such view entails domain-specific visual resources at the basis of VRT, which does not cohere with the fact that the maximum Leiter-3 score he obtained ('medium') was in CA, the least specifically visual subtest, which closely matches VRT. Furthermore, the visual processing difficulties revealed by ‘medium-low’ FG, FC, SO scores might explain why he failed in EIT (more dependent on visual working memory; Martins et al., 2016), which is easier than VRT for typical children. If recursion is a domain-general capacity, then why would it selectively fail to develop in language? The comparison between VRT and speech processing suggests that the motor and generative dimensions present in speech (that are absent in visual representation), in conjunction with the fast and transient character of the sequential auditory stimuli of speech, properties that have all been shown to be costly to manage in autism, could be responsible for the selective absence of linguistic recursion in AX.

References

Berwick, R.C. & Chomsky, N. (2016) *Why only us. Language and evolution.* MIT Press, Cambridge, MA.

Cook, D., Martins, M. D., & Villringer, A. (2018). How Domain-Specific is Merge. The Evolution of Language -Proceedings of the 12th International Conference, (1), 80–82. https://doi.org/10.12775/3991-1.016

Martins, M.D. (2012). Distinctive signatures of recursion. *Philosophical Transactions of The Royal Society B*., 367, 2055–2064. https://doi.org/10.1098/rstb.2012.0097

Martins, M. J., Fischmeister, F. P., Puig-Waldmüller, E., Oh, J., Geißler, A., Robinson, S., … Beisteiner, R. (2014a). Fractal image perception provides novel insights into hierarchical cognition. *NeuroImage*, 96, 300–308. https://doi.org/10.1016/j.neuroimage.2014.03.064

Martins, M. J. D., Krause, C., Neville, D. A., Pino, D., Villringer, A., & Obrig, H. (2019). Recursive hierarchical embedding in vision is impaired by posterior middle temporal gyrus lesions. *Brain*, 142(10), 3217–3229. https://doi.org/10.1093/brain/awz242

Martins, M. D., Laaha, S., Freiberger, E. M., Choi, S., & Fitch, W. T. (2014b). How children perceive fractals: Hierarchical self-similarity and cognitive development. *Cognition*, 133(1), 10–24. https://doi.org/10.1016/j.cognition.2014.05.010

Martins, M. D., Martins, I. P., & Fitch, W. T. (2016). A novel approach to investigate recursion and iteration in visual hierarchical processing. *Behavior Research Methods*, 48(4), 1421–1442. https://doi.org/10.3758/s13428-015-0657-1

Martins, M. D., Muršič, Z., Oh, J., & Fitch, W. T. (2015). Representing visual recursion does not require verbal or motor resources. *Cognitive Psychology*, 77, 20–41. https://doi.org/10.1016/j.cogpsych.2015.01.004

Roid, G.H. & Koch, C. (2017, 2nd.ed.) Leiter-3: Nonverbal cognitive and neuropsychological assessment. In R.S. McCallum (Ed.) *Handbook of nonverbal assessment.* Springer. doi10.1007/978-3-319-50604-3

LONG-RANGE SEQUENTIAL DEPENDENCIES PRECEDE SYNTACTICALLY-RICH VOCALIZATIONS IN HUMANS

TIM SAINBURG*[1], ANNA MAI[1], and TIMOTHY Q GENTNER[1]

*Corresponding Author: tsainbur@ucsd.edu
[1]Department of Psychology, University of California San Diego, La Jolla, USA

1. Introduction

As the distances between elements in human language sequences increase, the mutual information (MI) between them decays following a power law. This power-law relationship has been variously attributed to human language syntax (Lin & Tegmark, 2017), semantics (Alvarez-Lacalle, Dorow, Eckmann, & Moses, 2006), and discourse structure (Altmann, Cristadoro, & Degli Esposti, 2012). However, the vocalizations of numerous phylogenetically distant species including humpback whales and songbirds also demonstrate similar long-range statistical dependencies (Sainburg, Theilman, Thielk, & Gentner, 2019). These observations support the hypothesis that the long-range statistical dependencies found in human speech can occur independently of linguistic structure. To test this hypothesis, we computed MI over several child speech corpora (aged 6 months – 12 years) to determine whether long-range relationships are present in human vocalizations prior to the production of syntactically-rich speech.

2. Methods

We computed MI over words and phonemes for 9 corpora of natural speech from English speaking children included in the PhonBank and CHILDES repositories. Although data collection protocols differed across corpora, all data included here were transcripts of spontaneous speech from typically-developing children aged 6 months to 12 years. Transcripts were binned into five 6-month age groups and one age group from 3–12 years. Each transcript was analyzed either as sequences of words or as sequences of phonemes. Sequence distributions were treated independently across speakers to account for individual variation in lexical acquisition.

For each corpus, we calculated sequential MI over the elements of the sequence dataset (i.e., words or phonemes), and fit the MI decay using the methods outlined in Sainburg et al. (2019). The three models fit were (1) a power-law model, (2) an exponential model, and (3) a composite of models (1) and (2).

$$MI = ax^b + c \tag{1}$$

$$MI = ae^{xb} + c \tag{2}$$

$$MI = ae^{xb} + cx^d + f \tag{3}$$

where x represents the inter-element distance between units (e.g. words or phonemes). Context-free languages exhibit power-law decays in MI while Markovian processes exhibit exponential decays (Lin & Tegmark, 2017). The composite model has previously been shown to capture the sequential organization of human speech (Sainburg et al., 2019) where the exponential decay captures organization of speech elements within words and at short distance, and longer-range dependencies are captured by the power-law decay.

3. Results

MI decay for phonemes and in all age groups is best fit by the composite model, which includes a significant power law component. The same is true for words, with the exception of the 6-12 month age group (Figure 1), where decay is best fit by an exponential model. At 12 months, infants produce one-word utterances (Doupe & Kuhl, 1999). After 12 months, we observed a power-law at long distances and an exponential decay at shorter distances, consistent with adult language results. Our observation of the power law decay in infant speech contradicts the notion that it is generated solely by either human language syntax, semantics, or discourse structure.

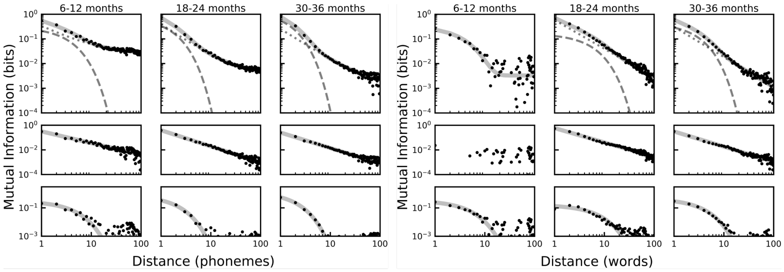


Figure 1.. MI decay over phonemes (left) and words (right) during development. Top row: Composite model fit; Middle row: Power law component fit; Bottom row: Exponential component fit.

4. Conclusion

Developmental emergence of long-range statistical dependencies in human speech precede the production of the linguistic structures hypothesized to generate them. This suggests that the long-range statistical organization of human speech is independent of human language syntax, semantics, or discourse structure. Instead, these dynamics may reflect a general property of biologically-generated sequences (Dawkins, 1976).

References

Altmann, E. G., Cristadoro, G., & Degli Esposti, M. (2012). On the origin of long-range correlations in texts. *Proceedings of the National Academy of Sciences*, *109*(29), 11582–11587.

Alvarez-Lacalle, E., Dorow, B., Eckmann, J.-P., & Moses, E. (2006). Hierarchical structures induce long-range dynamical correlations in written texts. *Proceedings of the National Academy of Sciences*, *103*(21), 7956–7961.

Dawkins, R. (1976). Hierarchical organisation: A candidate principle for ethology.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, *22*(1), 567–631.

Lin, H., & Tegmark, M. (2017). Critical behavior in physics and probabilistic formal languages. *Entropy*, *19*(7), 299.

Sainburg, T., Theilman, B., Thielk, M., & Gentner, T. Q. (2019). Parallels in the sequential organization of birdsong and human speech. *Nature communications*, *10*.

RATIONAL AFTER ALL: CHANGES IN PROBABILITY MATCHING BEHAVIOUR ACROSS TIME IN HUMANS AND MONKEYS

CARMEN SALDANA*[1,2,3], NICOLAS CLAIDIÈRE[4,5], JÖEL FAGOT[4,5], and KENNY SMITH[1]

*Corresponding Author: carmen.saldanagascon@uzh.ch
[1]Centre for Language Evolution, University of Edinburgh, Edinburgh, UK
[2]Center for the Interdisciplinary Study of Language Evolution, University of Zurich, Zurich, CH
[3]Department of Comparative Linguistics, University of Zurich, Zurich, CH
[4]Aix-Marseille University, CNRS, LPC, Marseille, FR
[5]Brain and Language Research Institute, Aix-Marseille University, Aix-en-Provence, FR

Probability matching strategies have long been thought to be characteristic of human performance in probability learning tasks in a variety of contexts, from decision making to language learning. Probability matching occurs when subjects given probabilistic input respond in a way that is proportional to the input probabilities. However, such behaviour is not optimal in a decision theoretic sense; the optimal decision strategy is to always select the variant with higher positive-outcome probability, known as maximising (or regularising, for linguistic tasks).

Propensity to probability match may differ across ages and species. While adults probability match in probability learning tasks, children tend to use maximising strategies instead: this difference across ages has been shown in linguistic as well as non-linguistic tasks (e.g., Derks & Paclisanu, 1967; Hudson Kam & Newport, 2005). However, more recent work comparing probability matching behaviour across linguistic and non-linguistic domains further suggests that adults are less likely to probability match in linguistic tasks (Ferdinand, Kirby, & Smith, 2019). Linguists have taken these differences across age groups and domains to suggest that regularisation of unconditioned linguistic variation over time might be driven by domain-specific biases as well as by domain-general biases not specific to humans. While very few studies have directly compared behavioural differences between primate species, existing studies suggest that monkeys ($N = 2$ to 8), unlike humans, adopt maximisation strategies (Parrish, Brosnan, Wilson, & Beran, 2014). These results suggest that probability matching behaviour might be restricted to adult humans, and perhaps most evident in non-linguistic domains. However, we lack direct robust evidence from the differences between human and non-human primate behaviour in simple decision making tasks, and a thorough exploration of how probability matching behaviour develops across time.

Here we present a series of experiments designed to directly compare probability matching behaviour across time in adult humans and Guinea baboons

(*Papio papio*) with a hitherto unmatched sample size (up to $N = 20$ baboons). The preregistered design and analysis plan for these experiments is accessible at `osf.io/qnm57`. We ran two experiments with different reward regimes. In Experiment 1 subjects were rewarded probabilistically (baboons, with food; humans, with money); in Experiment 2 subjects were always rewarded regardless of their response, lowering the cost of probability matching. In each experiment, we further manipulated the number of shapes (two or three) and reward probability (skewed or uniform). On each trial, subjects saw a set of coloured shapes (two or three shapes, randomly positioned on a computer screen) and were prompted to select one. Each shape lead to a reward according to the input ratio in the condition—70:30 (skewed, two shapes), 70:15:15 (skewed, three shapes), 50:50 (uniform, two shapes) or 33:33:33 (uniform, three shapes). If the subject selected the target shape for a given trial, they were rewarded. If the target shape was not touched, the subject proceeded to the next trial without reward (in Experiment 1) or to a recovery trial (in Experiment 2); in the recovery trial, the target image would be highlighted, subjects were prompted to select it and were then rewarded. Participants completed at least 240 trials. All factors were manipulated within-subjects for baboons ($N = 20$) and between-subjects for humans ($N = 160$).

In Experiment 1, where reward was probabilistic and the distribution of reward was skewed (i.e. 70:30 or 70:15:15), both species initially showed probability matching followed by a switch to maximising: in the first block of 60 trials, the selection of the shape with the highest reward probability was not significantly different from its reward probability, but there was a significant increase of maximising behaviour by block and final convergence to maximising behaviour after 240 trials. Crucially, we found probability matching behaviour in both species when the reward distribution was uniform. This difference in behaviour between skewed and uniform conditions suggests that maximising is not the default strategy but that both species are sensitive to the availability of maximising strategies.

In Experiment 2, in which reward was always available, we found that humans behaved as in Experiment 1 (i.e. probability matching then maximising with skewed and not with uniform distributions) but baboons responded randomly in all conditions (i.e. selecting all shapes with equal probability). These results suggest that humans maximised even when probability matching behaviour was not (monetarily) penalised, thus suggesting that maximising strategies in humans, unlike in baboons, are not uniquely driven by reward.

Our study provides evidence against the common assumption that humans probability match in simple decision making tasks and raises questions over the validity of conclusions in standard behavioural experiments, which our results suggest may simply have insufficient trials to show maximising or hide differences across time. It also casts doubt on the suggested domain-specific sources of maximising behaviour in linguistic tasks by providing evidence of shared maximising mechanisms in probability learning across primate species in non-linguistic tasks.

References

Derks, P. L., & Paclisanu, M. I. (1967). Simple strategies in binary prediction by children and adults. *Journal of Experimental Psychology*, *73*(2), 278.

Ferdinand, V., Kirby, S., & Smith, K. (2019). The cognitive roots of regularization in language. *Cognition*, *184*, 53–68.

Hudson Kam, C. L., & Newport, E. (2005). Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Lang. Learn. Dev.*, *1*(2), 151–195.

Parrish, A. E., Brosnan, S. F., Wilson, B. J., & Beran, M. J. (2014). Differential Responding by Rhesus Monkeys (Macaca mulatta) and Humans (Homo sapiens) to Variable Outcomes in the Assurance Game. *Anim. Behav. Cogn.*, *1*(3), 215–229.

LANGUAGE DISCRIMINATION BY LARGE-BILLED CROWS

SABRINA SCHALZ*[1], and EI-ICHI IZAWA[2]

*Corresponding Author: SS3903@live.mdx.ac.uk
[1]Department of Psychology, Middlesex University, London, UK
[2]Department of Psychology, Keio University, Tokyo, Japan

Language discrimination has previously been found in human infants, cotton-top tamarin monkeys, rats, and Java sparrows. This ability might also be relevant for the crow, a social passerine with extensive auditory perceptual skills living in close contact with humans. In this experiment we tested whether crows autonomously pay attention to spoken language, and whether they can discriminate a familiar, locally spoken language (Japanese) from an unfamiliar language (Dutch) without training. When presented with sentences spoken by multiple speakers, the crows showed significantly more responses to the Dutch than to the Japanese, which suggests that they discriminate two languages with distinctive linguistic features, and that they might also be more attentive to an unfamiliar language, Dutch, compared to a familiar one, Japanese. These results further extend the hypothesis that language discrimination is based on a general perceptual mechanism that predates the evolution of language and show that crows can voluntarily apply this mechanism to language outside of experimental set-ups.

1. Introduction

Although language as a whole is unique to humans, some cognitive abilities necessary for language are shared with other species. Ramus and colleagues (2000) showed that both human newborns and cotton-top tamarin monkeys (*Saguinus oedipus*) can discriminate between two languages from different rhythmic classes. They used Japanese (mora-timed) and Dutch (stress-timed) sentences spoken by four different female speakers per language and presented them to the infants and monkeys in a habituation/dishabituation design. In a second experiment, they synthetized these sentences to only include prosodic characteristics and removed lexical and phonetic information, as well as speaker variability (see Ramus & Mehler, 1999 for full description). They found that human infants failed to discriminate the natural stimuli, but successfully discriminated the synthetized stimuli containing only prosodic information. On the other hand, the tamarin monkeys were able to discriminate both types of stimuli despite speaker variability, although they performed better with the natural sentences than with the synthetized sentences. When presented with the synthetized stimuli played backwards, both the tamarins and the infants failed to discriminate the stimuli sets. The authors conclude from these observations that the ability to extract and process cues relevant for language discrimination likely preceded human speech, although humans and tamarins may use different cues for this task.

Taking into account a conceptual replication by Toro, Trobalon and Sebastián-Gallés (2003), this perceptual mechanism may date back even further. Their work with Long-Evans rats (*Rattus norvegicus*) using lever-press training with the original stimuli by Ramus et al. (2000) showed that just like human newborns and tamarin monkeys, rats are able to discriminate between synthetized stimuli of different rhythmical classes (stress-timed Dutch and mora-timed Japanese) when they are played forwards, but not backwards. Further research by Toro, Trobalon and Sebastián-Gallés (2005) showed that the rats were able to generalize previously learned prosodic cues to novel stimuli, and that they could only discriminate natural sentences produced by a single speaker, but not those produced by multiple speakers. It is indeed curious that all three species tested performed equally well with the synthetized stimuli, while there seems to be large variations in their performance with natural stimuli, possibly due to irrelevant information introduced by speaker variability.

The great number of parallels between birdsong and human language make passerines a well-suited model organism for biolinguistics (see Doupe & Kuhl, 1999), and their sensitivity to acoustic features can be extended to human language as well. To name just a few examples, Java sparrows can discriminate between prosodic patterns in Japanese spoken either with admiration or suspicion and generalize them to new sentences if the prosody remains familiar (Naoi, Watanabe, Maekawa & Hibiya, 2012), and discriminate English and Chinese sentences spoken by a bilingual speaker and generalize this discrimination to new sentences and a new speaker with training (Watanabe, Yamamoto & Uozumi, 2006). Zebra finches can discriminate between familiar and novel infant-directed songs and speech in English and Russian (Philmore, Fisk, Falk & Tsang, 2017), discriminate between trochees and iambs (Spierings, Hubert & ten Cate, 2017), use formant frequencies to discriminate the words wit and wet despite speaker variability (Ohms et al. 2009), and abstract prosodic patterns of human speech with prosodic stress on either the first or final syllable and generalize them to new stimuli (Spierings & ten Cate, 2014). Spierings and ten Cate (2014) conclude from this that “the sensitivity to prosodic cues is not linked to the possession of language and might have preceded language evolution, possibly originating from a pre-existing sensitivity to meaningful variation in pre-linguistic communicative sounds.”

Crows live in social groups or fission-fusion societies (Clayton & Emery, 2007), which requires them to vocally communicate with group members and identify conspecifics based on auditory cues. They can discriminate conspecifics based on their unique vocal signature (Kondo, Izawa & Watanabe, 2010), discriminate reliable and unreliable conspecifics based on their individual call (Wascher, Hillemann, Canestrari & Baglione, 2015), and recognize group members using audio-visual cues (Kondo, Izawa & Watanabe, 2012). In addition to conspecific calls, crows also discriminate between familiar and unfamiliar human voices, possibly because they often live in close contact with humans (Wascher, Szipl, Boeckle & Wilkinson, 2012).

Taking into account these extensive capabilities related to the auditory discrimination of individual calls and the self-motivated attention to voices of both conspecifics and heterospecifics, crows may also be attentive to linguistic features of languages spoken in their surroundings. The purpose of this experiment is therefore to examine whether crows autonomously pay sufficient attention to spoken language to discriminate a familiar from an unfamiliar language without prior training. Foregoing training and keeping the experimental set-up as naturalistic as possible has the advantage of showing more accurately the linguistic capabilities and the degree of attention to linguistic features wild urban crows living in close contact with humans display on their own. We used the same stimuli previously used in Ramus et al. (2000), and Toro et al. (2003, 2005) to allow for a more accurate comparison between the findings of this experiment and the previous language discrimination experiments with human infants, cotton-top tamarin monkeys, and rats. Such a comparison might highlight the analogies and heterogeneities between these species, and thereby provide further insights into the evolution of the mechanisms necessary for language discrimination.

2. Method

Eight large-billed crows (*Corvus macrorhynchos*; four males and four females) between the ages of two and four years were tested. One female crow was excluded from analysis due to lack of response. All subjects were caught in the prefectures Tokyo, Chiba, and Ibaragi with the permission from the Environmental Bureau of the Tokyo Metropolitan Government. The crows were housed in individual stainless steel-mesh home cages with a total of twenty-four crows in the room of the animal experimental facility at Keio University. Both caretakers and previous experimenters were native Japanese speakers. Before and after the experiment, they had access to food and water ad libitum.

The experiment was carried out in an outdoor aviary (W1.5 × D2.7 × H1.6 m). In the aviary, four perches were installed in the back, middle, front and the front-right corner approximately 1m above ground. A water basin was placed on the ground. Outside the aviary, a wireless loudspeaker (Sound Link Mini, Bose, USA) was placed next to the front-right corner for stimulus presentation, and a video camera (Handy-Cam HDR CX535, Sony, Japan) for recording the crows' behaviour was placed at 50 cm from the front end of the aviary.

We used twenty Dutch and twenty Japanese sentences as stimuli. They were all declarative, adult-directed, approximately 2.5 seconds long, and spoken by four female native speakers. After the habituation to the aviary on three consecutive days, the crows were tested for their responses to the Dutch and Japanese stimuli in a total of eight trials which were distributed over four days (i.e., two trials per day). Four crows received Dutch stimuli for the first four trials and Japanese stimuli for the last four trials, while the other three crows were assigned the opposite language order. Before the start of each trial, the crows were given 3–5 min for familiarization to the surroundings. Each trial consisted of four blocks of stimulus presentation with inter-block intervals of a 1–2-min silent period. Within each block, a set of ten sentences spoken by two different speakers was continuously presented twice in a random order. A 30 min silent period was inserted between the trials each day. The trial schedule including stimulus presentation was controlled by the programme PsychoPy 3 (Peirce, 2007). The sound level was set at a range between 70 and 80 dB across the perches.

According to the different behavioural responses to 1,000 Hz and 1,600 Hz tone stimuli between individual crows in a pilot experiment, either of two behaviours as response for each crow was measured during the stimulus presentation blocks from the video-recorded data: the amount of time they had their head lifted at least above the horizontal line, or the amount of time they sat on the right half of the front perch or on the perch in the front right corner close to the loudspeaker. Response times were coded in BORIS (Friard & Gamba, 2016). To normalize the response time to the stimuli varying slightly in their durations for each crow, we calculated the relative value of response time per 10 seconds to the total stimulus duration in each (see equation 1).

$$1)\ relative\ response\ time = \frac{total\ response\ time}{total\ stimulus\ duration} \times 10$$

The results were analysed using a generalized linear mixed model with an inverse Gaussian error distribution and a log link function. The model included the relative response time as an independent variable, the language as a fixed effect, and the individuals and the blocks within each trial as random effects. These analyses were performed using the free software R v.3.6.1 with the ‘lme4’ package (Bates, Maechler, Bolker & Walker, 2015). Significance of the independent variable was tested based on the Wald tests at the 0.05 level. Animal housing and the experimental protocols adhered to the guidelines of the Animal Care and Use Committee of Keio University.

3. Results

The model analysis produced a significant effect of the language variable with a negative coefficient for Japanese ($p < 0.001$, $t = -4.90$, $\beta \pm$ S.E. $= -0.39 \pm 0.08$; figure 1). This result suggests that the crows were significantly more attentive to the Dutch sentences than the Japanese sentences.

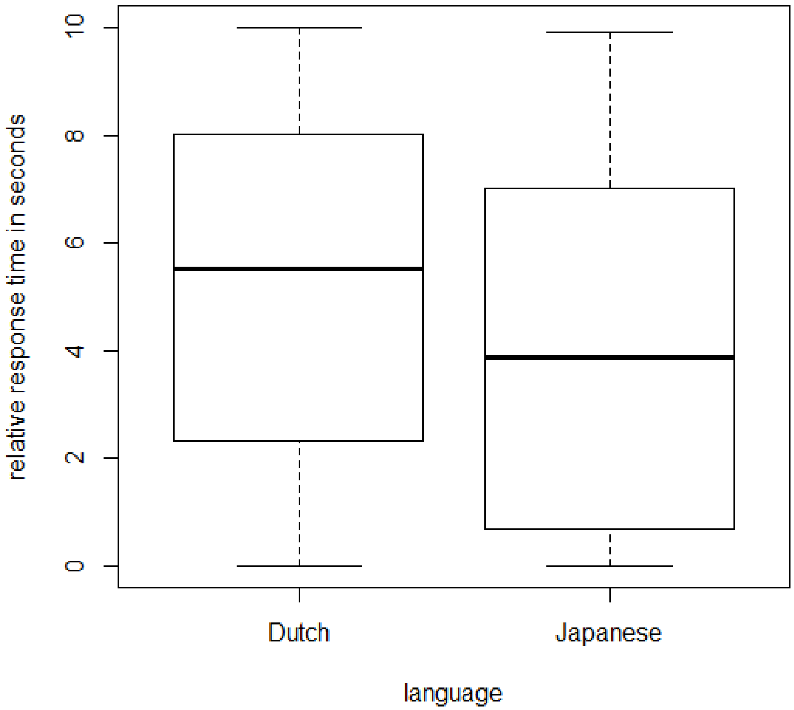


Figure 1. Relative response time per 10 seconds during the Dutch and the Japanese stimulus blocks. The crows showed more responses to the Dutch sentences than to Japanese ones.

At the individual level, five out of the seven crows clearly showed more responses to the Dutch stimuli that to the Japanese one (figure 2).

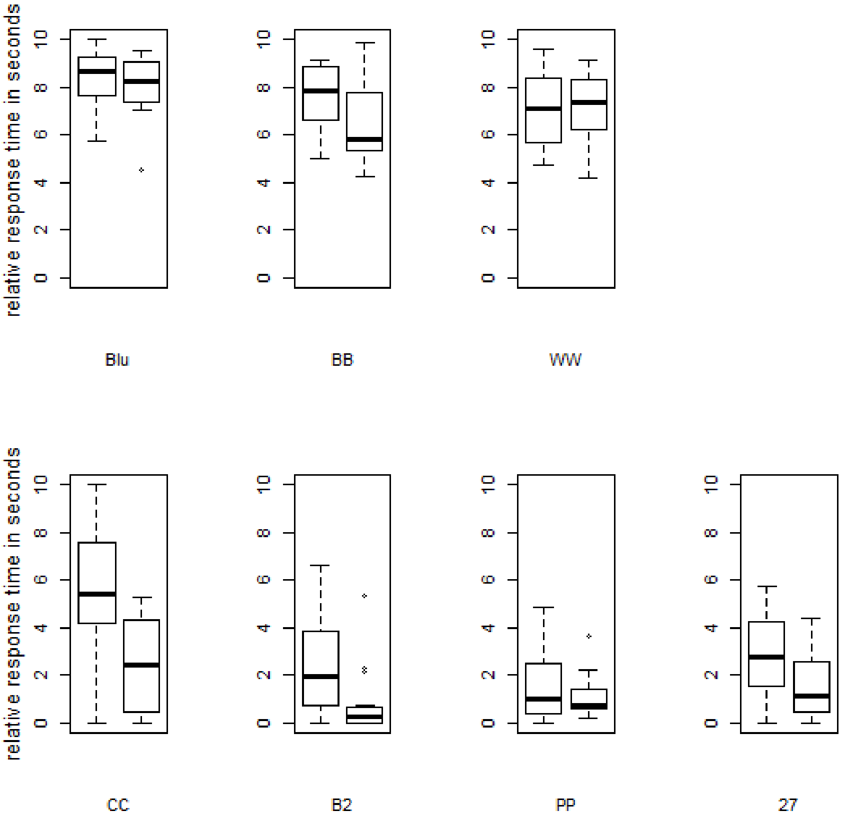


Figure 2. Individual relative response time per 10 seconds of each crow to the Dutch (left) and the Japanese (right) stimuli. The response behaviour for the crows in the first row was the amount of time their head was raised above the horizontal line, the response behaviour for the crows in the second row was the time they sat in the area next to the speaker.

4. Discussion

The results show that crows can discriminate between Dutch and Japanese sentences despite speaker variability and without prior training. A priori, crows should not be more interested in one language over the other. The initially higher attention to Dutch suggests that the crows were already familiar with Japanese before the experiment. Since all of them were caught in highly urbanized areas in Japan in and around Tokyo and were then in contact with Japanese experimenters and/or caretakers on a daily basis, it is safe to assumed that they were exposed to Japanese for their entire lives. This would support the hypothesis that crows actively listen to human speech of their own accord and without experimental set-ups to a degree that would enable them to identify and later recognize key features of Japanese independent of the individual speaker that distinguish it from other languages. Dutch, on the other hand, would likely be completely new to them and thus prompt them to pay more attention to it at first. This reaction would then be expected to gradually decline as they habituate more and more to it, and eventually their attention to Dutch should be equal to their attention to Japanese. The individual differences between the crows may be partially due to experience. “WW” and “Blu”, who were almost equally attentive to the two stimuli sets, are also the youngest crows at two and three years respectively and are considered juvenile, while the other crows are four years old. The shorter exposure to Japanese due to their young age might be the reason for their failure to discriminate it from Dutch, although further research is needed to verify this hypothesis.

Crows in urban areas such as Tokyo live in close contact with humans and speech would therefore be relevant to them, as it conveys information about the speaker’s identity and helps them determine whether they already know the specific person (Wascher et al. 2012) and whether that person might pose a threat. The perceptual abilities required for their extensive repertoire of vocalizations to communicate with conspecifics (Conner, 1985) and to discriminate group members based on their vocal signature (Kondo et al. 2010) may also be extended to the perception and categorization of human speech. Further experiments with crows from urban areas in other countries as well as crows from uninhabited areas are necessary to see whether the increased attention to the non-local language, or rather any language for crows from uninhabited areas, is consistently present.

These results stand in clear contrast with those obtained from human infants, who failed to discriminate the natural Dutch and Japanese sentences prior to the removal of non-prosodic information (Ramus et al. 2000). Speaker variability is likely the reason for this, as the rats successfully discriminated natural sentences spoken by only one speaker but failed when they were spoken by different speakers (Toro et al. 2005). The crows’ as well as the tamarins’ successful discrimination despite speaker variability points towards a more robust extraction of relevant linguistic features disregarding irrelevant information than that displayed by human infants and rats. Bird song and the vocalizations of New World monkeys show several similarities (see Snowdon, 1989), such as the repertoire of chirps and whistles used by cotton-top tamarin monkeys to convey different messages (Cleveland & Snowdon, 1982). Toro et al. (2005) argued that their experience with this type of vocalization, experience that rats do not have and infants have yet to gain, facilitates the discrimination task for the tamarin monkeys, which might also be the case with the crows. These results further support the previous findings in mammals and passerines that language discrimination is not a uniquely human ability and is instead based on a general perceptual mechanism that evolved prior to human language.

Taken together, the results obtained in this experiment show that crows living in close contact with humans are sufficiently attentive to spoken language out of intrinsic motivation to extract and recognize linguistic features distinguishing different languages from each other despite variation introduced by speaker variability. The crows' self-motivated attention to language could point towards an adaption to sharing their habitat with humans, as an increased attention to human vocalizations might provide information on danger, comparable to eavesdropping on heterospecific alarm calls observed in multiple species (e.g. Meise, Franks & Bro-Jørgensen, 2018). This attention to linguistic features may not be limited to language discrimination or the recognition of familiar voices. Further experiments are necessary to see which elements of language animals living in urban areas are also sensitive to, and whether there are any differences compared to individuals from rural areas.

Acknowledgements

We are very grateful to Frank Ramus for kindly supplying the original stimuli used in Ramus et al. (2000, 2005) and Toro et al. (2003, 2005). This experiment was carried out as part of the Thesis@Keio programme at Keio University. The study was financially supported by JSPS KAKENHI #17H02653 and Keio University Grant-in-Aid for Innovative Collaborative Research Projects #MKJ1905 to E.-I. I.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Clayton, N. S., & Emery, N. J. (2007). The social life of corvids. *Current Biology 17*(16), R652–656.

Cleveland, J., & Snowdon, C.T. (1982). The complex vocal repertoire of the adult cotton-top Tamarin (*Saguinus oedipus oedipus*). *Ethology 58*(3), 231–270.

Conner, R. N. (1985). Vocalizations of common ravens in Virginia. *The Condor 87*(3), 379–388.

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience 22*, 567–631.

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution 7*(11).

Kondo, N., Izawa, E.I., & Watanabe, S. (2010). Perceptual mechanism for vocal individual recognition in jungle crows (*Corvus macrorhynchos*): contact call signature and discrimination. *Behaviour 147*(8), 1051–1072.

Kondo, N., Izawa, E.I., & Watanabe, S. (2012). Crows cross-modally recognize group members but not non-group members. *Proceedings of the Royal Society B: Biological Sciences 279*(1735), 1937–1942.

Naoi, N., Watanabe, S., Maekawa, K., & Hibiya, J. (2012). Prosody discrimination by songbirds (*Padda oryzivora*). *PloS ONE 7*(10), e47446.

Meise, K., Franks, D.W., & Bro-Jørgensen, J. (2018). Multiple adaptive and non-adaptive processes determine responsiveness to heterospecific alarm calls in African savannah herbivores. *Proceedings of the Royal Society B: Biological Sciences 285*(1882), 20172676.

Ohms, V. R., Gill, A., van Heijningen, C. A. A., Beckers, G. J. L., & ten Cate, C. (2009). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society B: Biological Sciences* (1684), 1003–1009.

Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *Journal of Neuroscience Methods 162*(1-2), 8–13.

Philmore, L. S., Fisk, J., Falk, S., & Tsang, C. D. (2017). Songbirds as objective listeners: Zebra finches (*Taeniopygia guttata*) can discriminate infant-directed song and speech in two languages. *International Journal of Comparative Psychology 30*, uclapsych_ijcp_32722.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science 288*(5464), 349–351.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America 105*, 512–521.

Snowdon, C.T. (1989). Vocal communication in New World monkeys. *Journal of Human Evolution 18*(7), 611–633.

Spierings, M., Hubert, J., & ten Cate, C. (2017). Selective auditory grouping by zebra finches: testing the iambic-trochaic law. *Animal Cognition 20*(4), 665–675.

Spierings, M. J., & ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. *Proceedings of the Royal Society B: Biological Sciences 281*(1787), 20140480.

Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Animal Cognition 6*(2), 131–136.

Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2005). Effects of backward speech and speaker variability in language discrimination by rats. *Journal of Experimental Psychology. Animal Behaviour Processes 31*(1), 95–100.

Wascher, C. A. F., Hillemann, F., Canestrari, D., & Baglione, V. (2015). Carrion crows learn to discriminate between calls of reliable and unreliable conspecifics. *Animal Cognition 18*(5), 1181–1185.

Wascher, C. A. F., Szipl, G., Boeckle, M., & Wilkinson, A. (2012). You sound familiar: Carrion crows can differentiate between the calls of known and unknown heterospecifics. *Animal Cognition 15*(5), 1015–1019.

Watanabe, S., Yamamoto, E., & Uozumi, M. (2006). Language discrimination by Java sparrows. *Behavioural Processes 73*(1), 114–116.

MORPHOLOGICAL COMPLEXITY AND SIGN LANGUAGES: RETHINKING VERB DIRECTIONALITY IN EMERGING AND ESTABLISHED SIGN LANGUAGES

ANONYMOUS AUTHOR

*Corresponding Author: name@domain.com
[1]This Department, University X, City, Country

There has been a growing interest in the role emerging sign languages can play in our understanding of language evolution. In particular, work on home sign systems (e.g., Goldin-Meadow & Brentari, 2017), Nicaraguan Sign Language (e.g., Kocab et al., 2015) and Al Sayyid Bedouin Sign Language (e.g., Sandler, 2017) has provided a range of evidence about how grammar develops in new sign languages. Often this data is compared to current understanding of established sign languages, such as American Sign Language (e.g., Montemurro et al., 2019). However, existing assumptions about the grammar of established sign languages, such as British Sign Language (BSL) and its related variety Auslan (the majority sign language of Australia), are beginning to be challenged by new corpus data (e.g., Johnston, 2018; Fenlon et al., 2018). This paper will discuss concern that claims about language emergence need to be solidly grounded in a detailed understanding of the structure and use of both emerging and established sign languages. In particular, this paper will focus on the nature of verb directionality in sign languages. Indicating verbs in BSL, and other sign languages, can be directed towards locations in space associated with their arguments. This directionality has been widely analyzed as a form of person agreement marking (e.g., Lillo-Martin & Meier, 2011; Costello, 2016), although some scholars have rejected the agreement account (e.g., Corbett, 2006; Schembri et al., 2018). Indicating verbs have been considered an example of morphological 'complexification' in sign languages (e.g., Aronoff et al., 2005). The development of verb directionality has been studied in emerging sign languages, including Nicaraguan Sign Language, (Senghas & Coppola, 2001), as well as in artificial sign language studies (Motamedi et al., 2018). The data from both domains suggests that directionality develops over time from an initial stage in which arguments are represented by the signer's own body to one in which arguments are associated with abstract locations in space around the signer's body. Work drawing on spontaneous data has revealed, however, that the properties of indicating verbs in established sign languages are not as well understood as the literature might suggest. Fenlon et al. (2018) investigated a range of linguistic and social factors in 1,436 indicating verb tokens collected from a corpus of BSL, widely believed to be one of the oldest existing sign languages. Unlike previous claims based on elicited data (e.g., Morgan et al., 2006), the analysis shows that directionality in BSL is not obligatory (as had also been reported for Auslan, see de Beuzeville et al., 2009). Directionality is, in fact, variable, and conditioned by several linguistic factors, such as constructed action (a type of enactment used to represent referents), animacy, and co-reference. Unlike what is reported for emerging sign languages (Padden et al., 2010), the study did not find evidence of constructional change in progress in the BSL indicating verb system. Moreover, the corpus data show that the use of abstract spatial locations away from the signer's body is not the preferred strategy in BSL (Cormier et al., 2015). Only 9 tokens (4%) involving the use of abstract locations were identified out of 238 indicating verbs in third person to third person marking contexts. In fact, there is a strong preference for one of the arguments to be represented by the signer's body. This paper reports additional factors that may influence when a BSL sign is modified directionally for argument reference. Using the BSL Corpus indicating verb dataset, the influence of definiteness and variable argument noun phrase presence was investigated. This was based on the suggestion that definite referents might be established with more specific locations in the signing space compared to indefinite referents, as suggested by work on Catalan Sign Language (Barberà, 2016), and thus be more likely to trigger indicating verbs being directed towards these locations. The study also explores whether the presence of an explicitly expressed argument in the clause may be significant, something not previously investigated. Results suggest that definiteness is indeed a significant factor, with definite arguments more likely to trigger directionality in indicating verbs than indefinite arguments. Variable argument presence was also important with modification more likely in clauses with null argument expression. Thus, there are a complex range of factors that influence the use of verb directionality in a 'mature' sign language. This has important implications for an understanding of the emergence and development of morphological 'complexification' in sign languages, as the use of abstract space is clearly not obligatory, and – unlike what is suggested in the emerging and artificial sign language literature – actually appears to be disfavored. We need to be mindful of this when making claims about emerging versus established sign languages in discussions of sign language emergence, and its contribution to language evolution more generally.

References

Aronoff, M., Meir, I., & Sandler, W. (2005). The paradox of sign language morphology. *Language, 81*(2), 301-344.

Barbera, G. (2016). Indefiniteness and specificity marking in Catalan Sign Language (LSC). *Sign Language and Linguistics* 19:1, 1-36.

Corbett, G. (2006). *Agreement.* Cambridge: Cambridge University Press

Cormier, K., Fenlon, J. & Schembri, A. (2015). Indicating verbs in British Sign Language favour motivated use of space. *Open Linguistics* 1(1), 684-707.

Costello, B. (2016). *Language and modality: Effects of the use of space in the agreement system of Lengua de Signos Española (Spanish Sign Language).* LOT: Amsterdam.

de Beuzeville, L., Johnston, T., & Schembri, A. (2009). The use of space with indicating verbs in Australian Sign Language: A corpus-based investigation. *Sign Language and Linguistics* 12(1), 52-83.

Fenlon, J., Schembri, A., & Cormier, K. (2018). Modification of indicating verbs in British Sign Language: A corpus-based study. *Language, 94*(1), 84-118.

Goldin-Meadow, S. & Brentari, D. (2017). Gesture, sign, and language: The coming age of sign language and gesture studies. *Behavioral and Brain Sciences,* 40, E46.

Kocab, A., Pyers, J. & Senghas, A. (2015). Referential shift in Nicaraguan Sign Language: A transition from lexical to spatial devices. *Frontiers in Psychology,* 5: 1540.

Morgan, G., Barrière, I., & Woll, B. (2006). The influence of typology and modality on the acquisition of verb agreement morphology in British Sign Language. *First Language*, *26*(1), 19-43.

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J., & Kirby, S. (2018, April 14). The emergence of spatial modulation in artificial sign languages. https://doi.org/10.31234/osf.io/p6zy4

Padden, C., Meir, I., Sandler, W., & Aronoff, M. (2010). The grammar of space in two new sign languages. In D. Brentari (Ed.), *Sign Languages* (pp. 570-592). Cambridge: Cambridge University Press.

Sandler, W. (2016). What comes first in language emergence? In N. Enfield (Ed.), *Dependency in Language: On the Causal Ontology of Language Systems* (pp. 67-86). Berlin: Language Science Press. 67-86.

Schembri, A., Cormier, K., & Fenlon, J. (2018b). Indicating verbs as typologically unique constructions: Reconsidering verb 'agreement' in sign languages. *Glossa: A journal of general linguistics, 3*(1).

Senghas, A. & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science* 12(4), 323-328.

SIMULATING LEXICAL SEMANTIC CHANGE FROM SENSE-ANNOTATED DATA

DOMINIK SCHLECHTWEG* and SABINE SCHULTE IM WALDE

*dominik.schlechtweg@ims.uni-stuttgart.de
Institute for Natural Language Processing, University of Stuttgart, Germany

We present a novel procedure to simulate lexical semantic change from synchronic sense-annotated data, and demonstrate its usefulness for assessing lexical semantic change detection models. The induced dataset represents a stronger correspondence to empirically observed lexical semantic change than previous synthetic datasets, because it exploits the intimate relationship between synchronic polysemy and diachronic change. We publish the data and provide the first large-scale evaluation gold standard for LSC detection models.

1. Introduction

Evaluating Lexical Semantic Change (LSC) detection models is notoriously challenging. Existing testsets are flawed because they are too small to allow for generalizing over the results obtained on them. Artificial data, on the other hand, can be created in larger quantities, but typically relies on assumptions that may or may not be correct, such as the strength of semantic relatedness that old and new senses in LSC have. A clear advantage of artificial data is, however, that it allows the precise control of potentially influencing variables such as frequency and polysemy.

After spelling out the implicit assumptions of previous work, this paper presents a novel procedure to simulate lexical semantic change from synchronic sense-annotated data, which we consider more realistic than in earlier approaches. By splitting the synchronic data into two parts reflecting different sense frequency distributions for a word we simulate sense divergences. In a second stage, we define a graded and a binary notion of LSC based on differences between the obtained sense frequency distributions. These notions are then used to calculate the gold scores determining for each sense-annotated word the degree of change and whether senses were gained or lost. With the proposed definitions, we hope to provide a solid foundation for the basic concepts in the field of LSC detection.

2. Related Work

Most previous evaluations for LSC detection models rely on small amounts of empirically observed data, which was either hand-selected (Sagi et al., 2009; Jatowt & Duh, 2014; Hamilton et al., 2016a, 2016b; Frermann & Lapata, 2016; del Tredici & Fernández, 2017) or annotated by humans (Cook et al., 2014; Schlechtweg et al., 2017; Tahmasebi & Risse, 2017; Schlechtweg et al., 2018; Perrone et al., 2019). An alternative approach is synthetic evaluation, where pseudo-change is simulated by collapsing uses of different words (Cook & Stevenson, 2010; Kulkarni et al., 2015; Rosenfeld & Erk, 2018; Dubossarsky et al., 2019; Shoemark et al., 2019). This procedure is very similar to the creation of pseudo-polysemy in word sense disambiguation (Schütze, 1998; Pilehvar & Navigli, 2013): Changes in the senses a word w expresses are artificially created by copying a corpus and relabeling the uses of another word w' (known to express different senses) as uses of w in the copy. The word w is then guaranteed to express a different sense in the copy and the difference between the two corpora can be equated with artificial LSC.[1] This procedure mimics the changes in word senses occurring in empirical LSC, but requires assumptions about other factors. One such factor is the choice of words that should be collapsed and what their semantic relation should be. With the exception of Dubossarsky et al. (2019), all of the existing work collapses uses of words that have no semantic relation. This has the advantage that strong differences in the contextual distribution of words are created, which are more easily detected by computational models. However, there is plenty of evidence showing that LSC does not introduce random new senses of a word, but that new senses are very often semantically related to one of the old senses (Blank, 1997). Hence, the changes to the contextual distribution of a word in empirical LSC are often more subtle and hard to detect than the ones introduced in previous synthetic evaluations.

3. Simulation

Polysemy is the synchronic result of lexical semantic change (Blank, 1997; Bybee, 2015). Accordingly, the different senses a word may express today have been developed some time in the past by the word undergoing a process of LSC. We exploit this idea by using the modern (synchronic) senses of a polysemous word to simulate LSC, i.e., we reconstruct the diachronic process using its synchronic result. In this way, we (i) guarantee that the different senses used in the simulation are likely to be semantically related, as the different senses of a word are usually semantically related to each other (Fillmore & Atkins, 2000). And (ii) we simulate divergences for senses which have empirically been attached to the same word and are thus probable candidates to occur in empirical LSC.

3.1. *Corpus*

To simulate LSC in the above-described way we need sense-annotated data. We use **SemCor**, a sense-tagged corpus of English (Langone et al., 2004) which rep-

---

[1] Rosenfeld and Erk (2018) use a slightly different variation of this procedure.

resents a subset of the Brown Corpus (Francis & Kucera, 1979). SemCor contains 700,000 tokens, of which more than 200,000 are sense-annotated; the corpus is lemmatized and POS-tagged. Similar corpora in other languages (Bentivogli & Pianta, 2005; Henrich & Hinrichs, 2013, e.g.) would allow to easily extend our methodology to create evaluation data for further languages.

3.2. *Two concepts of LSC*

Graded. In order to simulate LSC, we first need a definition of what we mean by LSC. Previous research has not been explicit about the underlying concepts and often relied on an intuitive notion of *degree of LSC* (Hamilton et al., 2016b; Dubossarsky et al., 2017; Bamler & Mandt, 2017; Rudolph & Blei, 2018; Rosenfeld & Erk, 2018; Schlechtweg et al., 2018). A well-defined concept of graded LSC, however, should enable us to compare any two words over time and decide which of the words changed more. Hence, it should provide an answer to questions like: Did a word that lost a very frequent sense change more than a word that lost a very infrequent sense? And did a word that gained two senses change less than a word that lost three senses? The field is still lacking such a definition of the degree of LSC of a word.

Binary. This graded notion of LSC seems to diverge from the definition applied in historical linguistics, where LSC is typically not assumed to be graded, but binary (Blank, 1997, cf. p. 113). That is, either a word gained/lost a sense over time, or not, while in computational linguistics slight changes to the frequencies of different word senses are usually also considered as instances of LSC (hence the term *degree of LSC*).[2] This deviation is striking, as the most straightforward application of LSC detection models is their use to aid historical linguists (Hamilton et al., 2016b). Yet, the graded notion is applicable to related tasks, such as detecting the early stages of a meaning loss (as indicated by sense frequency decrease).

Table 1. Corpus sample for the noun *plant*.

This reduces the number of expensive **plant** shutdowns and startups.	**(s1)**
The pilot **plant** was equipped with a 3-hp. turbine aerator (Figure 2).	**(s1)**
Remove about half the branches from each **plant**, leaving only the strongest with the largest buds.	**(s2)**
"On the side toward the horizon – the southern hemisphere – it is spring; **plants** are being taught to grow".	**(s2)**
Can you share medical facilities and staff with neighboring **plants**??	**(s1)**

---

[2] Note for completeness that some work in computational linguistics also assumes a binary notion (Cook et al., 2014; Tahmasebi & Risse, 2017; Perrone et al., 2019; Shoemark et al., 2019).

3.2.1. *Sense Frequency Distributions*

We propose quantitative definitions of the two notions described above. The definitions are based on the concept of a Sense Frequency Distribution (SFD) (McCarthy et al., 2004; Lau et al., 2014, e.g.). A SFD encodes how often a word w occurs in each of its senses. Consider the word *plant*. In SemCor it occurs expressing either of two senses:

- **sense 1**: plant, works, industrial plant (buildings for carrying on industrial labor); "they built a large plant to manufacture automobiles"[3]
- **sense 2**: plant, flora, plant life (botany: a living organism lacking the power of locomotion)

This means that each use of *plant* in the corpus is assigned to one of these two senses, as we can see in the sample in Table 1: *plant* occurs three times in sense 1 and two times in sense 2. Hence, its SFD is provided as (3,2). Generally, given a set of w's uses U where each use is assigned to one of the senses in the sequence $S = (S_1, S_2, ..., S_i)$, w's SFD is defined by

$$T = (f(S_1), f(S_2), ..., f(S_i))$$

where $f(S_i)$ is the number of times any use from U was mapped to the ith sense in S.

Graded LSC. For two different sets of uses U_1 and U_2 the corresponding SFDs may differ. We will now define two measures quantifying the difference between any two same-sized SFDs, which will correspond to graded and binary LSC. Assume we have two SFDs T_1 and T_2 defined for the same word w and sense sequence S, but for two *different* use sets U_1 and U_2. We first normalize T_1 and T_2 to probability distributions P and Q by dividing each element by the total sum of the frequencies of all senses in the respective distribution. The degree of LSC of the word w is then defined as the Jensen-Shannon distance between the two normalized frequency distributions:

$$G(w) = JSD(P, Q)$$

where the Jensen-Shannon distance is the symmetrized square root of the Kullback-Leibler divergence (Lin, 1991; Donoso & Sanchez, 2017).[4] $G(w)$ ranges between 0 and 1 and is high if P and Q assign very different probabilities to the same senses.

---

[3] `https://wordnet.princeton.edu/`

[4] We prefer the Jensen-Shannon distance over Kullback-Leibler divergence, because the former is a true metric in contrast to the latter. This means that amongst other properties specific to metrics the Jensen-Shannon distance is symmetric, i.e., $JSD(X, Y) = JSD(Y, X)$.

Binary LSC. Binary LSC of the word w is then defined as

$$\begin{aligned} B(w) &= 1 \text{ if for some } i,\ P_i = 0.0 \text{ and } Q_i \geq k, \\ &\qquad \text{or vice versa.} \\ B(w) &= 0 \text{ else.} \end{aligned}$$

where X_i is the ith element in X and k is a probability threshold set to 0.1. That is, $B(w)$ will be 1 if there is a sense which has at least probability of 0.1 in either P or Q but a probability of 0.0 in the other (a meaning is gained or lost). If this is not the case, $B(w)$ will be 0. Note that $B(w)$ can be seen as an extreme special case of $G(w)$: if a sense is never assigned to any use in U_1, this sense will have probability 0.0 in P, which will cause $G(w)$ to be higher than for any other possible probability assignment to that sense. However, $G(w)$ will typically not be equal to 1.0 in such a case, as it is also sensitive to the probabilities of the other senses which may have changed only slightly or not at all.

3.3. *Algorithm*

In order to simulate LSC in SemCor we split it into two parts (C_1, C_2) and extract the corresponding SFDs (T_1, T_2) for each sense-annotated lemma w. From T_1 and T_2 the scores of above-defined gold notions of LSC follow directly. The splitting process has two steps: In step (i), we introduce strong changes for specific target lemmas. For this we sample all lemmas with a frequency between 100 and 1000 and split their sentences into senses, i.e., for each target lemma we randomly shuffle senses and split them at a random index into two subsets. Then we try to assign sentences with senses from the first subset to C_1 and sentences with senses from the second subset to C_2. This maximizes change, because senses will tend to have uses in only one of C_1, C_2. In step (ii), all remaining sentences are randomly shuffled, split in half and added to C_1, C_2 respectively. With this process a non-target lemma will tend to have a rather low change score.

Finally, we extract C_1 and C_2 sentence-wise to separate text files, and the LSC scores to a CSV file. For each sentence, all words (no punctuation) are extracted and replaced by their lemma if existent, else the lowercased token is extracted. Annotated phrases like *on_the_other_hand* are split into individual words to increase data size, i.e., *on the other hand*. The resulting corpora C_1 and C_2 have sizes with 0.34M and 0.36M tokens.

Consider Tables 2 and 3, showing the sample corpus from Table 1 split according to steps (i) and (ii) respectively. In Table 2, *plant* is treated as a target lemma and split according to step (i): sense 1 is assigned to C_2, while sense 2 is assigned to C_1. The resulting SFDs are $T_1 = (0, 2)$ and $T_2 = (3, 0)$, creating high change scores of $G(\textit{plant}) = 1.0$ and $B(\textit{plant}) = 1$. As the probability changes of the two senses (from 0.0 to 1.0 and vice versa) are the strongest possible, the graded change score is at its maximum. And as *plant* also loses and gains a sense from C_1 to C_2 it shows binary change.

Table 2. Sample corpus split for the target lemma *plant*. $T_1 = (0, 2)$, $T_2 = (3, 0)$, $G(w) = 1.0$ and $B(w) = 1$.

C_1	C_2
remove about half the branch from each **plant** leave only the strong with the largest bud **(s2)**	the pilot **plant** was equip with a 3 hp turbine aerator figure 2 **(s1)**
on the side toward the horizon the southern hemisphere it be spring **plant** are being teach to grow **(s2)**	this reduce the number of expensive **plant** shutdown and startup **(s1)**
	can you share medical facility and staff with neighboring **plant** **(s1)**

In Table 3, *plant* is treated as a non-target lemma and thus split according to step (ii): both senses are assigned uniformly to C_1, C_2. The resulting SFDs are $T_1 = (2, 1)$ and $T_2 = (1, 1)$, creating change scores of $G(\textit{plant}) = 0.14$ and $B(\textit{plant}) = 0$. The probabilities of each sense are relatively similar in the two corpora, which leads to a low graded change score and no binary change.

Table 3. Sample corpus split for the non-target lemma *plant*. $T_1 = (2, 1)$, $T_2 = (1, 1)$, $G(w) = 0.14$ and $B(w) = 0$.

C_1	C_2
remove about half the branch from each **plant** leave only the strong with the largest bud **(s2)**	on the side toward the horizon the southern hemisphere it be spring **plant** are being teach to grow **(s2)**
the pilot **plant** was equip with a 3 hp turbine aerator figure 2 **(s1)**	can you share medical facility and staff with neighboring **plant** **(s1)**
this reduce the number of expensive **plant** shutdown and startup **(s1)**	

3.4. *Testsets*

With the corpus split and the extracted change scores we have a large amount of evaluation data available. However, the change scores are subject to noise through non-annotated data. That is, non-annotated uses of words distort the sense frequency distributions on which the change scores are based. In order to minimize this noise we disregard each lemma w that has a relative frequency error $RE(w) \geq 0.5$, where

$$RE(w) = \frac{\#(w) - \#\textit{annotated}(w)}{\#\textit{annotated}(w)}$$

with $\#(w)$ being w's corpus frequency and $\#\textit{annotated}(w)$ the number of w's annotated uses. Hence, we allow at most a number of half of w's annotated uses to be added to these for w to be part of the testset. We additionally disregard any lemma with a lower frequency than 50 in either of C_1, C_2. This results in a testset containing 148 lemmas with different change scores. All the data is publicly available and can be used for LSC detection evaluation.[5]

---

[5] Find the data under: `https://www.ims.uni-stuttgart.de/data/lsc-simul`.

3.5. *Discussion*

The corpus splitting process described in Section 3.3 controls the degree of change introduced for a particular lemma. However, this process is not built on a particular theoretical model of LSC, i.e., a model of how the underlying sense probability distributions should change to be similar to true LSC. This also determines how much variables such as polysemy and frequency will correlate with simulated LSC in the resulting dataset. The way in which we chose to split the corpus implicitly introduces higher rates of change for more polysemous words, i.e., simulated LSC correlates with polysemy. Similarly, it introduces specific frequency patterns for strongly changing words, i.e., simulated LSC correlates with frequency change. Whether and to which degree this holds for true LSC is still debated, but it is clear that these variables strongly bias model predictions (Hellrich & Hahn, 2016; Dubossarsky et al., 2017). Thus, to make sure that model performances on our dataset do not stem from model biases towards these variables we recommend to report a polysemy and a frequency baseline. Only performances above these baselines can be safely attributed not to stem from model biases.

4. Model Evaluation

We give a short example of how to evaluate LSC detection models on our dataset. We train all vector space models with all alignment techniques from Schlechtweg et al. (2019) on C_1 and C_2 and apply two similarity measures (CD, LND) to the resulting representations to create change score predictions.[6] Then we use Spearman's ρ to compare the resulting rankings against the graded change scores and Average Precision (AP) to compare them against the binary change scores. The results are presented in Table 4.

Generally, models show rather weak performances on the testset. The performances for graded change are considerably lower than in Schlechtweg et al. (2019), which may be attributed to the much smaller corpus sizes and the resulting noise. As expected, the frequency and polysemy baselines show positive correlations with change scores. On average the models outperform the frequency baseline for graded and binary change, while the polysemy baseline is only outperformed for binary change. However, the best models always outperform both baselines. Thus, we can conclude that a range of models measure more than just polysemy or frequency change.

The best models are SGNS and SVD with OP and WI (see Table 4) as alignments. This is similar to previous results in that SGNS+OP+CD has outperformed other models and SVD showed generally high performance (Schlechtweg et al.,

---

[6] Find implementations at `https://github.com/Garrafao/LSCDetection`. Because of the very small corpus size we choose a large window size of $n = 10$ for all models, experiment with low dimensionalities $d = \{30, 100\}$ for SVD, RI and SGNS and train all SGNS with 30 epochs. We set $k = 5$ and $t = none$. The rest of parameters is set as in Schlechtweg et al. (2019).

Table 4. Best and mean ρ (Graded) and AP (Binary) scores across similarity measures (SIM). Scores are averaged over five iterations for models with a random component. The column 'model' gives the model with the best score. SGNS = Skip-Gram with Negative Sampling, CD = Cosine Distance, LND = Local Neighborhood Distance, SVD = Singular Value Decomposition, OP = Orthogonal Procrustes, WI = Word Injection, POLY = Polysemy Baseline, FREQ = Normalized Frequency Difference (NFD) Baseline, RAND = Approximate Random Baseline for Binary Classification.

Dataset	**Measure**	**Graded**			**Binary**		
		mean	*best*	*model*	*mean*	*best*	*model*
SEMCOR	SIM	0.159	**0.451**	SGNS+OP+CD	**0.182**	**0.376**	SVD+WI+LND
	POLY	**0.349**	0.349	-	0.151	0.151	-
	FREQ	0.120	0.120	-	0.110	0.110	-
	RAND	-	-	-	0.081	0.081	-

2019). The comparably high performance of WI alignment may be attributed to its strong noise-reducing effect on our small and thus noisy training corpora (Dubossarsky et al., 2019). A surprising observation is the performance of LND, as in the experiments of Schlechtweg et al. CD has constantly outperformed LND. This may be related to the difference between binary and graded change, as Schlechtweg et al. only evaluated on graded change.

5. Conclusion

We simulated lexical semantic change from synchronic sense-annotated data, introduced the first large-scale, synthetic gold standard for LSC detection and showed how to use it for evaluation. As part of our novel procedure, we provided quantitative definitions of various notions of LSC which implicitly underlie previous work; we thus provided a theoretical basis for artificial and empirical LSC detection evaluation. In the future, we will create further gold standards by exploiting sense-annotated data across languages and use our suggested LSC notions for the simulation of pseudo-change. We will also use the data to evaluate diachronic contextualized embeddings (Giulianelli, 2019; Hu et al., 2019).

The simulation procedure we proposed may also have applications in cognitive research on language evolution (Karjus et al., 2018; Nölle et al., 2018; Tinits et al., 2017) or more dialogue-oriented studies on meaning change (Pleyer, 2017), where it may be used to simulate the semantic development of words over generations or conversations. Similarly, different types of annotated data may be used to simulate specific types of LSC as e.g. literal and non-literal usages of words (Köper & Schulte im Walde, 2016), metaphoric uses (Köper & Schulte im Walde, 2017) or concrete and abstract uses (Naumann et al., 2018).

Acknowledgments

The first author was supported by the Konrad Adenauer Foundation and the CRETA center funded by the German Ministry for Education and Research (BMBF).

References

Bamler, R., & Mandt, S. (2017). Dynamic word embeddings. In D. Precup & Y. W. Teh (Eds.), *Proceedings of the 34th International Conference on Machine Learning* (Vol. 70, pp. 380–389). International Convention Centre, Sydney, Australia: PMLR.

Bentivogli, L., & Pianta, E. (2005). Exploiting parallel texts in the creation of multilingual semantically annotated resources: the MultiSemCor Corpus. *Natural Language Engineering*, *11*(3), 247261.

Blank, A. (1997). *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen.* Tübingen: Niemeyer.

Bybee, J. L. (2015). *Language change.* Cambridge, United Kingdom: Cambridge University Press.

Cook, P., Lau, J. H., McCarthy, D., & Baldwin, T. (2014). Novel word-sense identification. In *25th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers* (pp. 1624–1635). Dublin, Ireland.

Cook, P., & Stevenson, S. (2010). Automatically Identifying Changes in the Semantic Orientation of Words. In N. C. C. Chair, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, & D. Tapias (Eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10).* Valletta, Malta: European Language Resources Association (ELRA).

del Tredici, M., & Fernández, R. (2017). Semantic variation in online communities of practice. In *IWCS 2017 - 12th International Conference on Computational Semantics - Long papers.*

Donoso, G., & Sanchez, D. (2017). Dialectometric analysis of language variation in twitter. In *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects* (pp. 16–25). Valencia, Spain.

Dubossarsky, H., Hengchen, S., Tahmasebi, N., & Schlechtweg, D. (2019). Time-Out: Temporal Referencing for Robust Modeling of Lexical Semantic Change. In *Proceedings of the 57th annual meeting of the association for computational linguistics* (pp. 457–470). Florence, Italy: Association for Computational Linguistics.

Dubossarsky, H., Weinshall, D., & Grossman, E. (2017). Outta control: Laws of semantic change and inherent biases in word representation models. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 1147–1156). Copenhagen, Denmark.

Fillmore, C. J., & Atkins, B. T. (2000). Describing polysemy: The case of 'crawl'. In Y. Ravin & C. Leacock (Eds.), *Polysemy: Theoretical and computational approaches.* Oxford University Press.

Francis, W. N., & Kucera, H. (1979). *Brown corpus manual* (Tech. Rep.). Department of Linguistics, Brown University, Providence, Rhode Island, US.

Fremmann, L., & Lapata, M. (2016). A Bayesian model of diachronic meaning change. *Transactions of the Association for Computational Linguistics*, *4*, 31–45.

Giulianelli, M. (2019). *Lexical Semantic Change Analysis with Contextualised Word Representations.* Unpublished master's thesis, University of Amsterdam, Amsterdam.

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016a). Cultural shift or linguistic drift? Comparing two computational measures of semantic change. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 2116–2121). Austin, Texas.

Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016b). Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1489–1501). Berlin, Germany.

Hellrich, J., & Hahn, U. (2016). Bad Company—Neighborhoods in Neural Embedding Spaces Considered Harmful. In *Proceedings of COLING 2016* (pp. 2785–2796). Osaka, Japan.

Henrich, V., & Hinrichs, E. (2013). Extending the tüba-d/z treebank with germanet sense annotation. In I. Gurevych, C. Biemann, & T. Zesch (Eds.), *Language Processing and Knowledge in the Web* (pp. 89–96). Berlin, Heidelberg: Springer Berlin Heidelberg.

Hu, R., Li, S., & Liang, S. (2019). Diachronic sense modeling with deep contextualized word embeddings: An ecological view. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3899–3908). Florence, Italy: Association for Computational Linguistics.

Jatowt, A., & Duh, K. (2014). A framework for analyzing semantic change of words across time. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries* (p. 229238). IEEE Press.

Karjus, A., Blythe, R. A., Kirby, S., & Smith, K. (2018). Challenges in detecting evolutionary forces in language change using diachronic corpora. *CoRR*, *abs/1811.01275*.

Köper, M., & Schulte im Walde, S. (2016). Distinguishing literal and non-literal usage of German particle verbs. In *Proceedings of the 2016 conference of the north American chapter of the association for computational linguistics: Human language technologies* (pp. 353–362). San Diego, California: Association for Computational Linguistics.

Köper, M., & Schulte im Walde, S. (2017). Improving verb metaphor detection by propagating abstractness to words, phrases and individual senses. In *Proceedings of the 1st Workshop on Sense, Concept and Entity Representations and their Applications* (pp. 24–30). Valencia, Spain: Association for Computational Linguistics.

Kulkarni, V., Al-Rfou, R., Perozzi, B., & Skiena, S. (2015). Statistically significant detection of linguistic change. In *Proceedings of the 24th International Conference on World Wide Web, WWW* (pp. 625–635). Florence, Italy.

Langone, H., Haskell, B. R., & Miller, G. A. (2004). Annotating wordnet. In *Proceedings of the Workshop Frontiers in Corpus Annotation at HLT-NAACL.* Boston, MA, USA.

Lau, J. H., Cook, P., McCarthy, D., Gella, S., & Baldwin, T. (2014). Learning word sense distributions, detecting unattested senses and identifying novel senses using topic models. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 259–270). Baltimore, Maryland: Association for Computational Linguistics.

Lin, J. (1991). Lin jh.. divergence measures based on the shannon entropy. ieee trans inform theory 37: 145-151. *IEEE Transactions on Information Theory*, *37*, 145-I151.

McCarthy, D., Koeling, R., Weeds, J., & Carroll, J. (2004). Finding predominant word senses in untagged text. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics (ACL'04), Main Volume* (pp. 279–286). Barcelona, Spain.

Naumann, D., Frassinelli, D., & Schulte im Walde, S. (2018). Quantitative semantic variation in the contexts of concrete and abstract words. In *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics* (pp. 76–85). New Orleans, Louisiana: Association for Computational Linguistics.

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition*, *181*, 93–104.

Perrone, V., Palma, M., Hengchen, S., Vatri, A., Smith, J. Q., & McGillivray, B. (2019). GASC: Genre-aware semantic change for ancient Greek. In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change* (pp. 56–66). Florence, Italy: Association for Computational Linguistics.

Pilehvar, M. T., & Navigli, R. (2013). Paving the way to a large-scale pseudosense-annotated dataset. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 1100–1109). Atlanta, Georgia, USA.

Pleyer, M. (2017). Protolanguage and mechanisms of meaning construal in interaction. *Language Sciences*, *63*, 69–90.

Rosenfeld, A., & Erk, K. (2018). Deep neural models of semantic shift. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 474–484). New Orleans, Louisiana.

Rudolph, M. R., & Blei, D. M. (2018). Dynamic embeddings for language evolution. In *WWW 2018* (pp. 1003–1011). ACM.

Sagi, E., Kaufmann, S., & Clark, B. (2009). Semantic Density Analysis: Comparing Word Meaning Across Time and Phonetic Space. In *Proceedings of the Workshop on Geometrical Models of Natural Language Semantics* (pp. 104–111). Stroudsburg, PA, USA.

Schlechtweg, D., Eckmann, S., Santus, E., Schulte im Walde, S., & Hole, D. (2017). German in flux: Detecting metaphoric change via word entropy. In *Proceedings of the 21st Conference on Computational Natural Language Learning* (pp. 354–367). Vancouver, Canada.

Schlechtweg, D., Hätty, A., del Tredici, M., & Schulte im Walde, S. (2019). A Wind of Change: Detecting and Evaluating Lexical Semantic Change across Times and Domains. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 732–746). Florence, Italy: Association for Computational Linguistics.

Schlechtweg, D., Schulte im Walde, S., & Eckmann, S. (2018). Diachronic Usage Relatedness (DURel): A Framework for the Annotation of Lexical Semantic Change. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (p. 169-174). New Orleans, Louisiana.

Schütze, H. (1998). Automatic word sense discrimination. *Computational Linguistics*, *24*(1), 97–123.

Shoemark, P., Liza, F. F., Nguyen, D., Hale, S., & McGillivray, B. (2019). Room to Glo: A systematic comparison of semantic change detection approaches with word embeddings. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 66–76). Hong Kong, China: Association for Computational Linguistics.

Tahmasebi, N., & Risse, T. (2017). Finding individual word sense changes and their delay in appearance. In *Proceedings of the International Conference Recent Advances in Natural Language Processing* (pp. 741–749). Varna, Bulgaria.

Tinits, P., Nölle, J., & Hartmann, S. (2017). Usage context influences the evolution of overspecification in iterated learning. *Journal of Language Evolution*, *2*(2), 148-159.

LEARNING VS USE IN LANGUAGE EVOLUTION

THOM SCOTT-PHILLIPS[1,2], SEÁN G. ROBERTS[3], & ADAM JOWETT[3]

[1] Department of Cognitive Science, Central European University, Budapest, Hungary;
scott-phillipst@ceu.edu
[2] Department of Anthropology, Durham University, UK
[3] Department of Archaeology & Anthropology, University of Bristol, UK
[4] Department of Psychology, Royal Holloway University of London, UK

Communicatively efficient coding systems must balance two qualities that pull in partially different directions: the need to be expressive (useful in communication), and the need to keep complexity to a minimum. A prominent evolutionary hypothesis is that patterns of efficiency observed in natural languages are the product of two independent factors: (i) the need to be useful in communicative interaction shapes languages in the direction of sufficient expressivity; and (ii) the need to be acquirable by new generations of learners shapes languages in the direction of sufficient simplicity (sometimes called compressibility) (e.g. Kirby et al., 2015; Nölle et al., 2018; Raviv et al., 2018).

The most direct test of this hypothesis would: (1) systematically manipulate both independent variables (use & learning) such that the causal effects of each, and their combination, are contrasted and measured; (2) measure both dependent variables (usability & learnability) directly (rather than with proxies); (3) use as initial stimuli a language that had intermediate levels of both dependent variables (to allow appropriate change distinguishable from a ceiling effect). No existing experiment has all of these properties. (For example, the experiment in Kirby et al. (2015) manipulated learning (iterated chain or closed-group structure), but not use, and many studies use systematicity as a proxy for learnability.) This gap in the literature persists, we believe, due to the practical and methodological issues it entails.

Here we present such an experiment, and in doing so we introduce several novelties to experimental language evolution. We directly manipulate each factor across three conditions: learning-only (exposure to language, test); learn+use (exposure, communication game with a partner, test); use-only (communication game with a partner and access to an editable dictionary, test). The learnability of each output language was measured directly i.e. with naive learners learning them. The seed language had medium levels of learnability and expressivity (measured in the same way). A single transmission between the seed language and a new 'generation' was run four times in each condition. This allowed estimation of the general impact that each factor has on an evolving language.

Results show that only languages exposed to both factors, learning and communication, showed clear evidence of evolving to become both more learnable (*t*=8.88, *p*=.003) and more expressive (*t*=9.44, *p*=.003) (Figure 1). At the same time, our data also hint at the possibility that languages evolving only in the context of communication would still develop structure, albeit more slowly than languages that are also subject to a steady turnover of new learners.

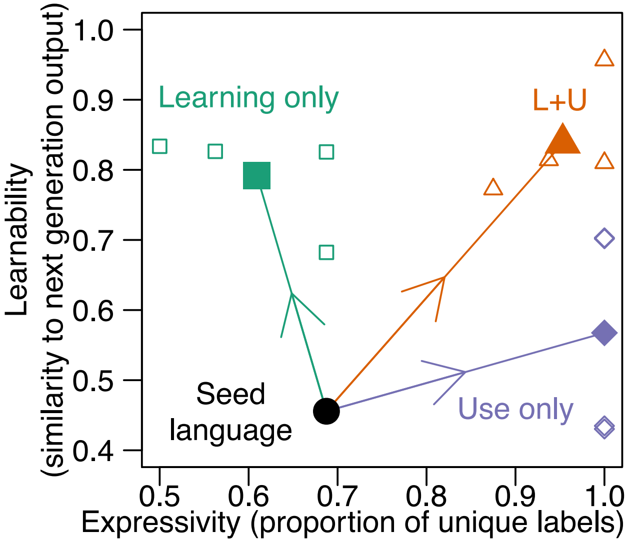


Figure 1. **Experimental results.** One seed language was exposed to each of three distinct experimental conditions. Unfilled shapes are individual data points; filled shapes are averages per condition. As can be seen, only languages in the learn+use condition gained significantly in both expressivity (*t*=8.88, df=3, *p*=.003) and learnability (*t*=9.44, df=3, *p*=.003). Full analyses of all conditions not presented only to preserve space.

We interpret these results alongside other data sources, of diverse types (e.g. Raviv et al., 2018; Lupyan & Dale, 2010; Senghas & Coppola, 2001; Bohn et al., in press), and also theoretical arguments from cognitive pragmatics about the cognitive mechanisms involved in communication (e.g. Scott-Phillips, 2015). Drawing things together, we suggest that language structure can evolve in response to communicative need alone, but this process may be accelerated by new learners. This proposal has the potential to reconcile prominent evolutionary approaches to language structure with the efficient communication hypothesis, which proposes that structure can arise only from the need to be useful in communication (e.g. Gibson et al., 2019). More detailed investigation of these ideas is an important frontier for language evolution.

Acknowledgements

TSP is supported by an ESRC Future Research Leaders grant (ES/K001647/1), and the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement no. 609819 (Somics project). SR is supported by a Leverhulme early career fellowship (ECF-2016-435).

References

Bohn, M., Kachel, G., & Tomasello, M. (in press). Young children spontaneously recreate core properties of language in a new modality. *PNAS.*

Gibson, E., Futrell, R., Piandadosi, S. T., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences, 23*(5), 389-407.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition, 141*, 87-102.

Lupyan, G., & Dale, R. (2010). Language structure is partly determined by social structure. *PLoS One, 5*(1), e8559.

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition, 181*, 93-104.

Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Compositional structure can emerge without generational transmission. *Cognition, 182*, 151-164.

Scott-Phillips, T. C. (2015). *Speaking Our Minds*. London: Palgrave MacMillan.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science, 12*(4), 323-328.

THE CO-EVOLUTION OF VALUE SYSTEMS AND INSTITUTIONS: A MICRO-SCALE MODEL OF DISSEMINATION OF COMMUNICATIVE VARIANTS

JOSE SEGOVIA-MARTIN*[1,2] and MONICA TAMARIZ[2]

*Corresponding Author: jose.segovia@e-campus.uab.cat
[1]Facultat de Filosofia i Lletres, Universitat Autònoma de Barcelona, Spain
[2]School of Social Sciences, Heriot-Watt University, United Kingdom

The idea that different value systems can lead to different distributions of variants in a population affecting cultural diversity has been explored in the past. For example, Axelrod's (1997) model of dissemination of culture was based on the assumption that people are more likely to interact with others who share the same cultural variants, and this in turn tends to increase the number of variants they share. Expanding on Axelrod (1997), researchers have found factors that affect the dynamics of the spread of variants: for example, globalization (Greig, 2002), mass media (Shibanai, Yasuno, & Ishiguro, 2001) and political institutions (Bhavnani, 2003). However, the complex network of interactions between value systems, population biases and institutional reinforcement has received comparatively little attention, in particular, when applied to micro-scale modelling of experimental designs of language evolution. How does the introduction of institutions that are sensitive to individuals' choices affect the maintenance of the diversity of variants (signals that represent one meaning) within the communicative system under enforced global connectivity? We develop an agent-based model to simulate micro-societies where we systematically manipulate:

1. Content bias (β): is a parameter identifying agents' sensitivity towards variants value (s). We examined values from 0 (no sensitivity) to 1 (full sensitivity) in steps of 0.1.

2. Value systems (S): it is a vector of floating-point numbers that correspond, for each agent, to the value (s) assigned to each cultural variant (σ) and indicates to what extent the variant is preferred over the other variants. We examined the evolution of the diversity of variants under two initial conditions of S: One takes all (OTA), where there is one preferred variant with value 1 and the rest with value 0; Pseudo random (PR), where each agent is assigned a value system S so that the value of each variant is a random floating point number N such that $0 \leq \mathrm{N} \leq 1$.

3. Institution (G). It is a vector that consists of the arithmetic mean of individual value systems S in the micro-society. G is weighted by two parameters:

institutional power (ε), that is, the capacity of the institution to effectively communicate its values to the agents (it takes values from 0 (no capacity) to 1 (full capacity) in steps of 0.1); and *conformity bias* (κ), which identifies agent's bias to conform to institutional values (it takes values from 0 (no conformity) to 1 (full conformity) in steps of 0.1.

In the initial state each agent i is randomly assigned a cultural variant σ_i and a value system S_i. At the beginning of each round, agents are paired randomly. Once agents are paired, at each round, they interact by presenting and observing one cultural variant. Within each pair, each agent in turn samples its history to produce a variant according to a probabilistic function that includes the parameters explained above. Then, both agents add both variants to their memories. After each interaction, agents' S, agents' probabilistic function of variant choice and G are updated according to the produced and observed variants, agents' biases, agents' record of variants, and the prior state of S and G.

Our results show that the maintenance of diversity in the communicative system is highly dependent on institutional performance. ε has the quantitative effect of decreasing diversity. However, this positive correlation is non-monotonic across β and κ. We show that institutional power reduces diversity both under OTA and under PR. In general, this effect is amplified by β. In both cases, especially under PR, the effect of institutions is stronger for intermediate values of κ (Figure 1). The latter point is important, for it means that in a context of high diversity of value systems, institutional intervention may cause stronger convergence on shared communicative variants when agents' behaviours are not extreme (not fully biased towards conformity). This is because full conformity with institutions might prevent agents from a faster alignment with their partners in their local interactions, slowing down convergence. Null institutional power is associated with higher diversity of variants. Our model can help to understand how institutions (e.g. prescriptivist vs. non prescriptivist institutions) direct the dynamics of communicative conventions, which might be useful for policy makers.

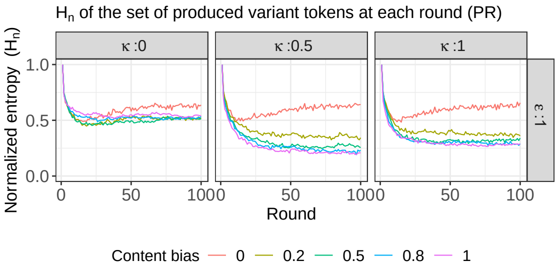


Figure 1. Diversity (measured as Normalized Shannon Entropy) over variants in the population. For each condition we performed 500 runs.

References

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, *41*(2), 203–226.

Bhavnani, R. (2003). Adaptive agents, political institutions and civic traditions in modern italy. *Journal of Artificial Societies and Social Simulation*, *6*(4).

Greig, J. M. (2002). The end of geography? globalization, communications, and culture in the international system. *Journal of Conflict Resolution*, *46*(2), 225–243.

Shibanai, Y., Yasuno, S., & Ishiguro, I. (2001). Effects of global information feedback on diversity: extensions to axelrod's adaptive culture model. *Journal of Conflict Resolution*, *45*(1), 80–96.

PHYLOGENETIC EXPLORATION OF LANGUAGE COMPLEXITY IN AUSTRONESIAN, BANTU, AND INDO-EUROPEAN LANGUAGE FAMILIES

OLENA SHCHERBAKOVA*[1], HEDVIG SKIRGÅRD[2], and SIMON J. GREENHILL [1,2]

*Corresponding Author: shcherbakova@shh.mpg.de
[1]Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena 07743, Germany
[2]Australian Research Council Center of Excellence for the Dynamics of Language, Australian National University, Canberra, ACT 0200, Australia

While language complexity has received attention from sociolinguistic, psycholinguistic, and computational perspectives, the processes of simplification and complexification over time remain challenging to examine and explain. One strand of research focuses on complexity 'tradeoffs' and 'local complexity' asking whether complexification in one grammatical domain necessitates simplification in another so that all languages are 'equi-complex' (Miestamo, 2009, Sinnemäki, 2008). The tradeoffs may or may not occur between different language systems, such as phonetics and morphology (Shosted, 2006), morphology and syntax (Dahl, 2009, Sinnemäki, 2008), morphosyntax and vocabulary size (Reali et al., 2018), and morphosyntax and semantics (Bisang, 2009).

However, most studies agree that local complexity across different domains varies between languages and there is little evidence for tradeoffs. Instead, there appear to be evidence of considerable variation in the causes and effects of complexity. For example, measures of complexity do not respond to extra-linguistic factors in the same way; Sinnemäki and Di Garbo (2018) show that verbal inflectional synthesis negatively correlates with the population size of speech communities, while nominal complexity of grammatical gender does not. Linguistic paradigms appear to be more challenging to transmit within larger communities (Nettle, 2012, Reali et al., 2018) as they are harder to acquire by L2 learners, resulting in creoles being paradigmatically simpler (Good, 2012).

Here we aim to quantify and explore these issues using a large global dataset, Grambank (Skirgård et al. in review), comprised of 195 grammatical features from over 1,800 languages. To measure nominal and verbal complexity, we selected features signaling the adherence of the languages to the principle of distinctiveness (Sinnemäki, 2009). While the presence of these features facilitates comprehension for hearers providing additional grammatical information, they simultaneously contribute to the complexity from the speakers' perspective impeding the ease of production and requiring more efforts to articulate (Mufwene, 2012). Our metric of nominal complexity encompasses the presence of marking for such categories like number, gender, possessiveness, and case. Conversely, our metric of verbal complexity accounts for marking of arguments, overt signaling of tenses, aspect, and other markers on verbs. Finally, we measure the paradigmatic complexity of function words based on the distinctions existing between articles, pronouns, demonstratives, classifiers, and adpositions, which so far have been overlooked in most complexity studies.

We use these metrics to quantify language complexity along three different axes: nominal, verbal, and paradigmatic complexity. We apply cutting-edge phylogenetic methods to model change in these measures of complexity and explore how they have diversified, changed, and traded-off over time in the Austronesian, Bantu, and Indo-European language families. We then undertake a path analysis while controlling for phylogenetic non-independence to disentangle the evolutionary relationships between the measured complexity types *(*van der Bijl, 2018).

Our results suggest that, first, the paradigmatic complexity of function words does not interact with the nominal and verbal complexity. In contrast, nominal and verbal complexity display a weak positive correlation in Bantu and Indo-European languages. However, the correlation in the Indo-European family disappears once creole languages are removed from the analysis. Furthermore, the investigation of causal mechanisms via path analysis on a global dataset indicates that nominal complexity tends to influence verbal complexity. Our results suggest that the processes of simplification and complexification show no overall global trends but prove to be dependent on the domain and language family.

References

Bisang, W. (2009). On the evolution of complexity: sometimes less is more in East and mainland Southeast Asia. In G. Sampson, D. Gil and P. Trudgill (Eds.), *Language Complexity as an Evolving Variable* (pp. 34-49). Oxford: Oxford University Press.

Dahl, Ö. (2009). Testing the assumption of complexity invariance: The case of Elfdalian and Swedish. In G. Sampson, D. Gil and P. Trudgill (Eds.), *Language Complexity as an Evolving Variable* (pp. 50-63). Oxford: Oxford University Press.

Good, J. (2015). Paradigmatic complexity in pidgins and creoles. *Word Structure, 8*(2), 184-227.

Miestamo, M. (2009). Implicational hierarchies and grammatical complexity. In G. Sampson, D. Gil and P. Trudgill (Eds.), *Language Complexity as an Evolving Variable* (pp. 81-97). Oxford: Oxford University Press.

Mufwene, S. S. (2012). The emergence of complexity in language: An evolutionary perspective. In A. Massip-Bonet and A. Bastardas-Boada (Eds.), Complexity perspectives on language, communication and society (pp. 197-218). Springer Verlag.

Nettle, D. (2012). Social scale and structural complexity in human languages. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1597), 1829-1836.

Reali, F., Chater N., & Christiansen, M. H. (2018). Simpler grammar, larger vocabulary: How population size affects language. *Proceedings of the Royal Society B: Biological Sciences, 285*(1871), 20172586.

Shosted, R. K. (2006). Correlating complexity: A typological approach. *Linguistic Typology, 10,* 1-40.

Sinnemäki, K. (2008). Complexity trade-offs in core argument marking. In M. Miestamo, K. Sinnemäki and F. Karlsson (Eds.), *Language complexity: Typology,contact, change* (pp. 67-88). Amsterdam, Philadelphia: Benjamins.

Sinnemäki, K. (2009). Complexity in core argument marking and population size. In G. Sampson, D. Gil and P. Trudgill (Eds.), *Language Complexity as an Evolving Variable* (pp. 126-140). Oxford: Oxford University Press.

Sinnemäki, K., & Di Garbo, F. (2018). Language structures may adapt to the sociolinguistic environment, but it matters what and how you count: a typological study of verbal and nominal complexity. *Frontiers in psychology, 9,* 1141.

Skirgård et al. (in review). Grambank, Jena: Max Planck Institute for the Science of Human History.

Trudgill, P. (2011). *Sociolinguistic typology: social determinants of linguistic complexity*. Oxford: Oxford University Press.

van der Bijl, Wouter. (2018). phylopath: Easy phylogenetic path analysis in R. PeerJ 6:e4718 https://doi.org/10.7717/peerj.4718

THE COMMUNICATIVE DEMANDS OF TOOLMAKING SKILL TRANSMISSION

DOR SHILTON[*1]

[*]Corresponding Author: dorshilt@mail.tau.ac.il
[1] Cohn Institute for the History and Philosophy of Science and Ideas, Tel-Aviv University, Tel-Aviv, Israel

Hypotheses about the co-evolution of language and toolmaking consider both the cognitive demands of production and the communicative demands of social transmission. Regarding the latter, it has been suggested that increasingly complex toolmaking technologies required and selected for more complex communication technologies (Gärdenfors and Högberg 2017, Goren-Inbar et al. 2018, Laland 2018; for a different approach, see Tennie 2017). I evaluate this hypothesis considering Dor's (2015) theory of language and Donald's (1991) theory of mimesis. According to Dor (2015), the distinguishing feature of language is that it is an instructive communication technology rather than an experiential one; it aims not to show an experience (as one does, for example, with pointing and pantomime), but rather to provide the receiver with instructions on how to imagine it. Its unique function is therefore the instruction of imagination. Donald (1991), as well as other researchers, has suggested a mimetic stage in human evolution, in which communication was intentional, representational and multi modal, but lacked compositionality and arbitrariness. This experiential communication technology would have been highly valuable for interactions confined to the here-and-now, including the ones involved in the social transmission of skills. Consequently, I argue that the unique function of language - the instruction of imagination - is not necessary for toolmaking skill transmission, and that mimetic communication, its likely precursor, is sufficient. To demonstrate this, I review evidence from both ethnographic and experimental studies.

Hunter-gatherer ethnographies suggest that the social learning of skills relies mainly on observation, experimentation, participation, and play (e.g. McDonald

2007, Hewlett 2016, Lew-Levy et al. 2017). Explicit instruction is rare, a fact that might be related to the emphasis on personal autonomy common to forager societies. More elaborate use language, in conversations and storytelling, focuses primarily on social norms. Studies of traditional stone cultures (Hampton 1999, Stout 2002) describe how the social transmission of toolmaking skill is facilitated through group activities with simple, context-bound interactions embedded in the here and now. Feedback consists mostly of gestures and short utterances (e.g. ‘do it here’, ‘don’t do that’) that are easily emulated by mimetic communication.

Experimental studies comparing gestural and verbal teaching of toolmaking skills also demonstrate that gestural communication is sufficient, and most have found that language does not improve transmission (Onhuma et al. 1997, Putt et al. 2014, Morgan et al. 2015, Lombao et al. 2017, Cataldo et al. 2018). I therefore conclude that mimesis would have been sufficient for the social transmission of toolmaking skills. The ethnographic evidence suggests the same is true for other hominin skills like hunting and foraging. Finally, I would suggest that in terms of cultural demands, it wasn’t teaching per se that drove language evolution, but rather the structuring of social norms through complaints, gossip and storytelling.

References

Cataldo, D. M., Migliano, A. B., & Vinicius, L. (2018). Speech, stone toolmaking and the evolution of language. *PloS one*, 13(1), e0191071.

Donald, M. (1991) *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: Harvard University Press.

Dor, D. (2015) *The Instruction of Imagination: Language as a Social Communication Technology*. Oxford: Oxford University Press.

Gärdenfors P., Högberg A. (2017) ‘ The Archaeology of Teaching and the Evolution of Homo Docens’, Current Anthropology , 58/2: 188–208.

Goren-Inbar N. et al. (2018) The Acheulian Site of Gesher Benot Ya‘Aqov Volume IV: The Lithic Assemblages . Cham, Switzerland: Springer.

Hampton, O. W. (1999) *Culture of Stone: Sacred and Profane Uses of Stone among the Dani*. College Station, TX: Texas A&M University Press.

Hewlett, B. S. (2016). Social Learning and Innovation in Hunter-Gatherers. In Terashima, H., Hewlett, B. S. (Eds), *Social Learning and Innovation in Contemporary Hunter-Gatherers* (pp. 1–15). Tokyo: Springer.

Laland K. N. (2018) Darwin's Unfinished Symphony: How Culture Made the Human Mind. Princeton, NJ: Princeton University Press.

Lew-Levy, S., Reckin, R., Lavi, N., Cristóbal-Azkarate, J., & Ellis-Davies, K. (2017). How do hunter-gatherer children learn subsistence skills? *Human Nature*, 28(4), 367-394.

Lombao, D., Guardiola, M., & Mosquera, M. (2017). Teaching to make stone tools: new experimental evidence supporting a technological hypothesis for the origins of language. *Scientific reports*, 7(1), 14394Stout D. (2002) Skill and Cognition in Stone Tool Production: An Ethnographic Case Study from Irian Jaya. *Current Anthropology*, 43/5: 693–722.

MacDonald, K. (2007). Cross-cultural comparison of learning in human hunting. *Human Nature*, 18(4), 386-402.

Morgan, T. J., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., ... & Whiten, A. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature communications*, 6, 6029.

Ohnuma, K., Aoki, K., Akazawa, T. (1997) Transmission of Tool-Making through Verbal and Non-Verbal Communication: Preliminary Experiments in Levallois Flake Production. *Anthropological Science*, 105/3: 159–68.

Putt, S. S., Woods, A. D., Franciscus, R. G. (2014) The Role of Verbal Interaction during Experimental Bifacial Stone Tool Manufacture. *Lithic Technology*, 39/2: 96–112.

Tennie, C., Premo, L. S., Braun, D. R., & McPherron, S. P. (2017). Resetting the null hypothesis: early stone tools and cultural transmission. Current Anthropology, 58, 664-672.

DO LIFE HISTORIES SHAPE VOCAL PRODUCTION LEARNING DIFFERENCES IN BIRDS? AN EXPLORATORY APPROACH

SARA SILVENTE I FONT*[1], PEDRO TIAGO MARTINS[2,3], and CEDRIC BOECKX[2,3,4]

*Corresponding Author: sara.silvente@uclouvain.be
[1] Psychological Sciences Research Institute, University of Louvain, Louvain-la-Neuve, Belgium
[2]Section of General Linguistics, University of Barcelona, Barcelona, Spain
[3]University of Barcelona Institute of Complex Systems, Barcelona, Spain
[4]Catalan Institute for Research and Advanced Studies, Barcelona, Spain

1. Introduction

Vocal production learning (VPL) is the ability to change vocal output as a result of experience (e.g. auditory, be it through modification of spectral and/or temporal aspects of vocalizations or completely novel calls (see (Janik & Slater, 2000) for an influential definition). Despite the structural differences between the human cortex and the avian pallium, similar neural pathways have been proposed for VPL (Fitch, 2017; Jarvis, 2007), involving a direct forebrain (primary motor cortex/arcopallium) projection to the phonatory muscles (larynx/syrinx). Higher encephalization has been associated with complex behavioral traits (Isler & Schaik, 2009), including VPL (Liu, Wada, Jarvis, & Nottebohm, 2013). Charvet and Striedter (2011) proposed that the telenchephalic expasion undergone by VPL birds is the result of delayed and protracted neurogenesis in this brain region, with a major effect in some song nuclei during song learning (Liu et al., 2013). In turn, this form of post-hatching maturation of the telencephalin would be promoted by an altricial developmental mode (Charvet & Striedter, 2011). In summary, prolonged brain maturation is likely to lead to an extended developmental period that would benefit the development of complex behavioral traits including VPL. Defining VPL is a non-trivial issue, which we put aside in the present work, which is of an exploratory nature. With this in mind, we opt for a bottom-up approach in which we aim to explore whether differences between birds considered to be VPL and non-VPL naturally emerge from other data. For example, life histories are affected by body mass, such that larger species mature more slowly (Minias & Podlaszczuk, 2017). Avian species that learn their vocalizations tend to have relatively small body sizes (Liu et al., 2013). These species would be expected to group closer together by virtue of their small body size. The goals of the present work are the following: i) to explore whther groups reflecting divisions based on VPL/Non-VPL profiles emerge from avian life history traits (i.e, development trajectories) and two biometric measures, brain mass and body mass; ii) to test whether differences in (adult) relative brain mass (measure here by the ration of brain-to-body mass) are related to life history traits and contribute to shaping the differences between the groups obtained above.

2. Methods

Data. Six developmental variables representative of life history were used: incubation operiod, fledging age, period of post-fledging parental care, age of sexual maturity for males and for females, and maximum lifespan. Absolute body size and brain mass were used as biometric variables, as well as the brain-to-body mass ratio (as an approximate measure of relative brain mass). **Sample.** These data were extracted for 1498 species (34 orders) from the literature and public databases (Hoyo et al., 1992; Iwaniuk & Nelson, 2003; Striedter & Charvet, 2008; Myhrvold et al., 2015). After exclusing species with missing or unclear data, the final sample consists of 179 species (96 thought to be vocal learners) with several orders and families represented. **Analyses.** ***clustering*** In order to explore how VPL and non-VPL birds grouped as a function of their life histoies and biometric measures, we ran a hierarchical clustering analysis. Since developmental ctraits are correlated with one another, we specified Spearman correlation coefficients as the distance metric between clusters. Two hierarchical clusters were built by using the average and Ward's methods (average method: 0.674, Ward's method: 0.657). Non-hierarchical methods were also used to determine whether the groups obtained above emerged naturally from the data (e.g. k-metoid clustering using Partitioning Around Medoids (PAM). Different statistics were calculated over a range of 2 to 7 possuble k values. ***Permutation***. In order to extend the descriptive results, we tested the different groups obtained for significant differences using a permutation test (n $=$ 19999, no replacement). When statistically significant differences were found, Monto Carlo-based permutation was used (n = 9999) to assess the effect of random shuffling. Since avian life history traits are correlated with relative brain size, we tested whether the groups obtained in the hierarchical clustering analysis differed in life history traits when controlling for brain-to-body mass ratio (n = 10000).

3. Results & discussion

Our analyses yielded the following 3 meaningful clusters: VL (vocal leaners, 30 species), NVL (non-VL, 31 species), and a mix of both (MIX, 118 species, 66 vocal learners) (see supplementary materials). VL emerged as the most consistent group throughout our different analyses. Our findings suggest body size to range too widely to be a predictor of bird vocal VPL abilities. We found brain-to-body mass ratio to be related to all life history traits, but it has a significant effect in distinguishing the three groups when related to period of incubation and age of sexual maturity. We discuss the implications of our results for the study of VPL in birds, and perhaps other species, as well as some cautionary notes.

References

Charvet, C. J., & Striedter, G. F. (2011). Developmental Modes and Developmental Mechanisms can Channel Brain Evolution. *Frontiers in Neuroanatomy*, *5*.

Fitch, W. T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*, *24*(1), 3-33.

Hoyo, J. del, Elliott, A., Sargatal, J., Christie, D., International Council for Bird Preservation, & BirdLife International. (1992). *Handbook of the birds of the world.* Barcelona: Lynx [for] ICBP. (OCLC: 225496301)

Isler, K., & Schaik, C. P. van. (2009). The expensive brain: A framework for explaining evolutionary changes in brain size. *Journal of Human Evolution*, *57*(4), 392–400.

Iwaniuk, A. N., & Nelson, J. E. (2003). Developmental differences are correlated with relative brain size in birds: A comparative analysis. *Canadian Journal of Zoology*, *81*(12), 1913-1928.

Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60*(1), 1–11.

Jarvis, E. D. (2007). Neural systems for vocal learning in birds and humans: A synopsis. *Journal of Ornithology*, *148*(1), 35–44.

Liu, W.-c., Wada, K., Jarvis, E. D., & Nottebohm, F. (2013). Rudimentary substrates for vocal learning in a suboscine. *Nature Communications*, *4*, 2082.

Minias, P., & Podlaszczuk, P. (2017). Longevity is associated with relative brain size in birds. *Ecology and Evolution*, *7*(10), 3558–3566.

Myhrvold, N. P., Baldridge, E., Chan, B., Sivam, D., Freeman, D. L., & Ernest, S. K. M. (2015). An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles: *Ecological Archives* E096-269. *Ecology*, *96*(11), 3109-000.

Striedter, G. F., & Charvet, C. J. (2008). Developmental origins of species differences in telencephalon and tectum size: Morphometric comparisons between a parakeet (*Melopsittacus undulatus*) and a quail (*Colinus virgianus*). *The Journal of Comparative Neurology*, *507*(5), 1663-1675.

DOMINANCE STYLE AND VOCAL COMMUNICATION IN NON-HUMAN PRIMATES

KATIE SLOCOMBE[*1], EITHNE KAVANAGH1[1,], SALLY STREET[2], FELIX O. ANGWELA[3], THORE J. BERGMAN[4], MARYJKA B. BLASZCZYK[5], LAURA M. BOLT[6], MARGARITA BRISEÑO-JARAMILLO[7], MICHELLE BROWN[8], CHLOE CHEN-KRAUS[9], ZANNA CLAY[2], CAMILLE COYE[10], MELISSA EMERY THOMPSON[11], ALEJANDRO ESTRADA[7], BARBARA FRUTH[12], CLAUDIA FICHTEL[13], MARCO GAMBA[14], CRISTINA GIACOMA[14], KIRSTY E. GRAHAM[1], SAMANTHA GREEN[15], CYRIL GRUETER[15], SHREEJATA GUPTA[1], MORGAN L. GUSTISON[16], LINDSEY HAGBERG[17], DANIELA HEDWIG[18], KATHARINE M. JACK[19], PETER M. KAPPELER[13], GILLIAN KING-BAILEY[19], BARBORA KUBĚNOVÁ[20], ALBAN LEMASSON[10], DAVID MACGREGOR INGLIS[21], ZARIN MACHANDA[22], ANDREW MACINTOSH[20], BONAVENTURA MAJOLO[23], SOPHIE MARSHALL[1], STEPHANIE MERCIER[24], JÉRÔME MICHELETTA[25], MARTIN MULLER[11], HUGH NOTMAN[26], KARIM OUATTARA[27], JULIA OSTNERA[13], MARY SM PAVELKA[28], LOUISE R. PECKRE[13] , MEGAN PETERSDORF[29], FREDY QUINTERO[24], GABRIEL RAMOS- FERNÁNDEZ[30], MARTHA M. ROBBINS[31], ROBERTA SALMI[32], ISAAC SCHAMBERG[17], OLIVER SCHÜLKE[13], STUART SEMPLE[21], JOAN B. SILK[33], J. ROBERTO SOSA-LOPÉZ[34], VALERIA TORTI[14], DARIA VALENTE[14], RAFFAELLA VENTURA[35], ERICA VAN DE WAAL[36], ANNA H. WEYHER[37], CLAUDIA WILKE[1], RICHARD WRANGHAM[17], CHRISTOPHER YOUNG[38], ANNA ZANOLI[14], KLAUS ZUBERBÜHLER[24], ADRIANO LAMEIRA[39]

[*]Corresponding Author: ks553@york.ac.uk
[1]Department of Psychology, University of York, York, UK
[2] Department of Anthropology, Durham University, Durham, UK
[3] School of Agricultural and Environmental Sciences, Mountains of the Moon University, Fort Portal, Uganda
[4] Department of Psychology, University of Michigan, Ann Arbor, USA
[5] Department of Anthropology, University of Texas, Austin, USA
[6]Department of Anthropology, University of Waterloo, Canada,
[7]Instituto de Biologia, Universidad Nacional Autonoma de México, Mexico
[8] Department of Anthropology, University of California, Santa Barbara, USA
[9] Department of Anthropology, Yale University, New Haven, USA
[10] Human and Animal Ethology (EthoS), University of Rennes, Rennes, France
[11]Department of Anthropology, University of New Mexico, Albuquerque, USA
[12] School of Biological and Environmental Science, Liverpool John Moores University, UK,
[13]German Primate Center, Leibniz Institute for Primate Research, Göttingen, Germany

[14]Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy
[15] School of Human Sciences, University of Western Australia, Australia
[16]Department of Integrative Biology, University of Texas, Austin, USA
[17]Department of Human Evolutionary Biology, Harvard University, Cambridge, USA
[18]Center for Conservation Bioacoustics, Cornell University, Ithaca, USA
[19]Department of Anthropology, Tulana University, New Orleans, USA
[20]Primate Research Institute, Kyoto University, Japan
[21]Department of Life Sciences, University of Roehampton, London, UK
[22]Department of Anthropology, Tufts University, Medford, USA
[23]School of Psychology, University of Lincoln, Lincoln, UK
[24]Department of Comparative Cognition, University of Neuchâtel, Switzerland
[25]Department of Psychology, University of Portsmouth, Portsmouth, UK
[26]Department of Anthropology, Athabasca University, Canada
[27]Centre Suisse de Recherches Scientifiques en Côte d’Ivoire, Côte d'Ivoire
[28]Department of Anthropology and Archaeology, University of Calgary, Canada
[29]Department of Anthropology, New York University, New York, USA
[30]Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Mexico City, Mexico
[31]Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Leipizg, Germany
[32]Department of Anthropology, University of Georgia, Athens, USA
[33]School of Human Evolution and Social Change, Arizona State University, Tempe, USA
[34]Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional Unidad Oaxaca, Instituto Politécnico Nacional, Oaxaca, Mexico
[35]School of Social and Health Sciences, University of Abertay, Dundee, UK
[36]Department of Ecology and Evolution, University of Lausanne,Switzerland
[37]Department of Anthropology, University of Massachusetts, Amherst, USA
[38]Mammal Research Institute, University of Pretoria, Republic of South Africa
[39]Department of Psychology, University of Warwick, Coventry, UK

Understanding the variables that shape the use and evolution of vocal communication in non-human primates can inform understanding of how language evolved. Social complexity might drive communicative complexity (Freeberg et al. 2012). Dominance style (the strictness with which the dominance hierarchy is enforced; ranging from ‘despotic’ to ‘tolerant’; de Waal and Luttrell, 1989) is an important, but often overlooked, measure of social complexity and its relationship with vocal communication is largely unknown. As the outcomes of social interactions in more tolerant societies are more uncertain (Dobson, 2012), we predicted that more tolerant individuals and species would have a greater need for more frequent and more diverse vocal signals to negotiate their social interactions. Here, we provide evidence that dominance style is associated with vocal usage and repertoires at both individual and phylogenetic levels in primates.

At the inter-individual level, we considered given tolerance and received tolerance separately as there are different reasons for expecting dominant and subordinate individuals within tolerant relationships to communicate more frequently than those in despotic relationships. Considering these two measures separately should also allow us to infer whether tolerance puts pressure on dominant individuals to communicate more, or whether a more tolerant social environment relaxes constraints on subordinate individuals' communication. For our four behavioural dominance style variables (aggression symmetry, counter aggression, aggression intensity and grooming symmetry) we were able to obtain given and received tolerance measures for individuals by including only interactions with lower-ranking partners, or higher-ranking partners, respectively, in their calculations. We predicted that both i) given and ii) received tolerance versions of each variable would be associated with a higher rate of vocalising. At the interspecific level, we calculated the four dominance style variables per species, and combined them into a 'dominance style index'. We predicted that this index would be associated with three aspects of vocal repertoires, all of which were obtained from previous literature. Using Bayesian analyses on these observational data from 111 wild groups of 26 species, we show that more tolerant individuals vocalise at a higher rate, but more despotic species have a wider range of hierarchy-related vocalisations in their repertoires. We found little evidence that tolerance received from higher-ranking partners is related to vocal rate, or that more tolerant species have larger vocal repertoires in terms of overall repertoire size or number of social vocalisations. Our findings indicate that tolerance is related to vocal usage more strongly as a result of increased pressure for more tolerant individuals to communicate more, than alleviation of constraints on communication for lower-ranking individuals. Taken together, our findings indicate that dominance style is a valuable social variable for understanding vocal usage and evolution in primates.

References

De Waal, F. B., & Luttrell, L. M. (1989). Toward a comparative socioecology of the genus Macaca: different dominance styles in rhesus and stumptail monkeys. *American Journal of Primatology*, *19(2)*, 83-109.

Freeberg, T. M., Dunbar, R. I., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Phil. Trans. R. Soc. B*, *367*, 1785–1801

Dobson, S. D. (2012). Coevolution of facial expression and social tolerance in macaques. American Journal of Primatology, 74(3), 229-235

VARIATION IN MINDREADING “ON THE GROUND”: COMPARING PATTERNS OF MENTAL STATE TALK IN TWO SOCIETIES

ANDREW MARCUS SMITH*[1], H. CLARK BARRETT[1], and ERICA A. CARTMILL[1]

*Corresponding Author: andrewmsmith@ucla.edu
[1]Anthropology Department, University of California Los Angeles, Los Angeles, United States of America

A broad literature in pragmatics has underscored the importance of mindreading to the function of human language (Sperber & Wilson, 1986). Because a given utterance is likely to be semantically underdetermined, successful communication requires individuals to read the minds of their interlocutors and ascertain the intended meaning of the utterance. In brief, much of the expressive power of language is derived from users’ ability to read the minds of their interlocutors. These claims accord with the vast psychological literature on mindreading that has suggested mindreading is likely to be an early-developing and ontogenetically canalized core cognitive domain underlying much of human social behavior (e.g., Castelli et al., 2000; Senju et al., 2009). Such findings have bolstered claims suggesting there are few, if any, meaningful differences in adult mindreading ability across both individuals and cultures. However, these same findings often equivocate competence and performance (cf. Wu & Keysar, 2007 for an exception). While human beings may have an implicit and species-typical capacity, or competence, to impute others’ mental states, the way in which such imputations structure social behavior, or performance, may differ across populations according to prevailing socioecological conditions. An emerging anthropological literature lends credence to this claim (Robbins & Rumsey, 2008). Given the documented importance of mental-state talk for children’s early sociocognitive development (Ruffman, Slade, & Crowe, 2002), understanding whether patterns of mental-state talk are universal or variable may provide indirect evidence about the evolved architecture of the mindreading and language systems. Despite the centrality of these questions, few quantitative studies of mental-state talk across societies have been conducted.

To begin filling this lacuna, we measured mental state language in a small-scale society with implicit norms against attributing mental states to others. A simple response-elicitation task was administered in Achuar to bilingual Achuar / Spanish speakers in a small-scale, hunter-horticulturalist society in Amazonian Ecuador (N=40) as well as in English to a sample of American undergraduates (N=26). Participants were shown a set of nine silent videos ranging in length from 40 to 70 seconds long. Eight of these videos depicted interactions between two or more individuals in a narrative arc derived from one of the following fitness domains – Cooperation, Dangerous Animal, Dominance, Infidelity, Mate Guarding, Norm Violation, Prestige, and Sickness. To the extent that these domains have borne on fitness across human evolutionary history, there ought not to be sizable differences across populations in the extent to which these videos are interpretable by viewers across cultures. Moreover, the criteria according to which the videos were designed minimize reliance on culturally specific information. The ninth video was structured as a classic False Belief task performed in a naturalistic setting. Given the extensive literature demonstrating that four-year-old children succeed in attributing mental states to agents in the False Belief task, we hoped this video would serve as a standard against which to test the efficacy of the new stimuli in eliciting mental-state attributions.

After viewing each video, participants were asked to describe what had happened in the scene. Given previous ethnographic observation about Achuar speech, Achuar participants were expected to describe scenes using less mental-state language than American participants. Verbal responses were transcribed and coded according to a scheme adapted from Castelli et al. (2000) and Ruffman et al. (2002). Counts of words attributing affective states, perceptions, desires, and epistemic states to characters in the video were obtained. Word counts were scaled to control for differences in description length. Hierarchical Poisson Regression models of word counts were run with culture and video type as fixed factors and participant as a random factor.

Pilot studies using a similar methodology, but different stimuli found that American participants used words attributing perception, affective states, and epistemic states to characters significantly more often than Shuar participants, an ethnic group whose language is closely related to Achuar ($p<0.001$). Additionally, Shuar participants used desire words more frequently ($p<0.001$). Analyses of the present study conducted with Achuar participants are expected to demonstrate a similar pattern. Data on cross-cultural differences in mental state talk may illuminate the extent to which human language is constrained or free to vary in the conceptualization and communication of mental states.

References

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. Monographs of the society for research in child development, i-174.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage*, 12(3), 314-325.

Clark, H.H. (1996). *Using language.* Cambridge University Press.

Corballis, M.C., & Lea, S.E. (1999). *The descent of mind: Psychological perspectives on hominid evolution.* Oxford University Press.

Robbins, J., & Rumsey, A. (2008). Introduction: Cultural and linguistic anthropology and the opacity of other minds. *Anthropological Quarterly, 81*(2), 407-420.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child development, 73*(3), 734-751.

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: an absence of spontaneous theory of mind in Asperger syndrome. *Science, 325*(5942), 883-885.

Sperber, D., & Wilson, D. (1986). Relevance: Communication and cognition (Vol. 142). Cambridge, MA: Harvard University Press.

Tylén, K., Weed, E., Wallentin, M., Roepstorff, A., & Frith, C.D. (2010). Language as a tool for interacting minds. Mind & Language, 25(1), 3-29.

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. Psychological science, 18(7), 600-606.

DID PARTIAL HEARING LOSS SHAPE AUSTRALIAN PHONOLOGY? AN EMPIRICAL TEST USING SPOKEN ITERATED LEARNING

KENNY SMITH*[1] and DAN DEDIU[2]

*Corresponding Author: Kenny.Smith@ed.ac.uk
[1]Centre for Language Evolution, University of Edinburgh, Edinburgh, UK
[2]Laboratoire Dynamique Du Langage, Université Lumière Lyon 2, Lyon, France

Iterated learning experiments (where participants are trained on a miniature linguistic system then reproduce that system, with the reproduction being used as training for subsequent participants) form an important tool in evolutionary linguistics, and have been used to show how fundamental structural properties of language evolve through transmission (e.g. Kirby et al., 2008, 2015; Beckner et al., 2017). While iterated learning experiments have been run in acoustic and visual modalities (e.g. Verhoef et al., 2014; Motamedi et al., 2019), the foundational results on the evolution of compositional structure come from studies using written (typed) text. Since the neuro-congnitive mechanisms involved in reading and writing are not identical to those of spoken language (e.g. Huettig et al., 2018), replicating results in the spoken modality is necessary.

Moreover, moving to the spoken modality allows us to test hypothesised links between extra-linguistic biases affecting speech perception/production and the structure of spoken languages (e.g. Everett et al., 2016; Blasi et al., 2019). Here we test the intriguing proposal (Butcher, 2006) that certain typologically unusual properties of the consonant inventories of the Australian aboriginal languages – no *voicing* contrasts, few *manner*, but many *place* of articulation contrasts (but see Gasser & Bowern, 2014) – are due to the high prevalence, throughout history, of chronic middle-ear infections (*chronic otitis media* or COM) in aboriginal Australian children (World Health Organization, 2004). The proposal is that the ensuing hearing loss in a large proportion of the speaker population, mainly affecting the low and high frequencies where cues to voicing and manner contrasts reside, forced the Australian languages to adapt, dropping hard-to-hear cues and capitalising instead on the intact frequency regions (Butcher, 2006).

We report three experiments: *Experiment 1* replicates the "classic" iterated learning results from Kirby et al.'s (2008) Experiment 1 in the spoken modality. Participants were trained (using spoken stimuli) on a miniature language for describing patterned shapes and then asked to reproduce (in speech) the labels for those shapes. The initial participant in each chain was trained on randomly-generated labels, but subsequent participants were trained on the labels produced by the previous participant. The spoken languages evolve over generations, becoming more accurately learned and developing systematic underspecification.

Experiment 2 tests two *dyad-based conditions* (following Kirby et al., 2015): two individuals are trained (simultaneously but in separate experiment booths) on the same target language, then interact using speech (via audio streaming between booths), taking turns to label objects for their partner or to select objects based on their partner's label. In the *Chains condition* the set of labels produced by one member of the dyad is used to train a new pair of participants at the next generation; in the *Closed Group condition* one dyad plays for many rounds, with no naive participants being introduced. Our results broadly replicate those of Kirby et al. (2015): structure gradually increases over generations in the Chains condition but is relatively flat across rounds in Closed Groups. However, we found far more variability in the spoken modality than in the written modality, including lower alignment between the members of interacting dyads.

Finally, *Experiment 3* focuses on testing Butcher's (2006) hypothesis: we contrast the Closed Groups from Experiment 2 (the *Unfiltered* condition) with a new set of Closed Groups (the *Filtered* condition) where a real-time band-pass filter (filtering out frequencies below 400Hz and above 4KHz, simulating the after-effects of COM) was applied to all audio during training and interaction. While the languages which developed in *Filtered* and *Unfiltered* conditions looked broadly similar in terms of communicative accuracy, stability and structure, there were subtle differences in the consonant inventories used. While there was no difference in entropy of *manner* of articulation ($\beta = 0.02 \pm 0.03, p = .44$) or *place-of-articulation* entropy (i.e. no evidence for use a wider range of places of articulation under auditory filtering: $\beta = 0.02 \pm 0.03, p = .58$), Filtered dyads did have lower *voicing* entropy (i.e. a tendency for either voiced *or* voiceless consonants: $\beta = -0.03 \pm 0.013, p = .03$). This difference develop rapidly during communicative interaction and is in line with Butcher's hypothesis, providing preliminary empirical support for a role of COM in shaping some typologically interesting aspects of the phonology of Australian languages.

In sum, we show that iterated learning can be extended to the spoken modality, broadly replicating the previous findings and allowing us to test hypotheses concerning phonetic and phonological diversity, suggesting that the effects of COM might shape some aspects of phonology even within a single generation.

Acknowledgements

Thanks to Hanna Jarvinen, Amira Saouri and Rachel Kindellan for conducting data collection. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no 681942). DD was supported by an IDEXLYON (16-IDEX-0005) Fellowship grant (2018-2021).

References

Beckner, C., Pierrehumbert, J. B., & Hay, J. (2017). The emergence of linguistic structure in an online iterated learning task. *Journal of Language Evolution*, *2*, 160–176.

Blasi, D. E., Moran, S., Moisik, S. R., Widmer, P., Dediu, D., & Bickel, B. (2019). Human sound systems are shaped by post-neolithic changes in bite configuration. *Science*, *363*(6432).

Butcher, A. (2006). Australian aboriginal languages: Consonant salient phonologies and the ‘place-of-articulation imperative. In J. M. Harrington & M. Tabain (Eds.), *Speech production: models, phonetic processes and techniques* (pp. 187–210). New York, NY: Psychology Press.

Everett, C., Blasí, D. E., & Roberts, S. G. (2016). Language evolution and climate: the case of desiccation and tone. *Journal of Language Evolution*, *1*(1), 33-46.

Gasser, E., & Bowern, C. (2014). Revisiting Phonotactic Generalizations in Australian Languages. In *Proceedings of the Annual Meetings on Phonology* (Vol. 1).

Huettig, F., Kolinsky, R., & Lachmann, T. (2018). The culturally co-opted brain: how literacy affects the human mind. *Language, Cognition and Neuroscience*, *33*(3), 275-277.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences, USA*, *105*, 10681-10686.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87-102.

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J., & Kirby, S. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition*, *192*, 103964.

Verhoef, T., Kirby, S., & Boer, B. de. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, *43*, 57–68.

World Health Organization. (2004). *Chronic suppurative otitis media: burden of illness and management options.* Geneva, Switzerland: World Health Organization.

CEREBELLAR FORWARD MODELS AND THEIR ROLE IN AUDITORY PROCESSING

STEFANIE STURM*[1,2] and IRIA SANMIGUEL[1,2]

*Corresponding Author: stefaniesturm@ub.edu
[1]Brainlab – Cognitive Neuroscience Research Group, University of Barcelona, Spain
[2]Department of Clinical Psychology and Psychobiology, Institute of Neurosciences, University of Barcelona, Spain

Evidence is accumulating of the cerebellum as a rapidly evolved brain structure supporting a previously underappreciated variety of cognitive processes. This goes far beyond its more established role in motor control is a promising subject of study for cognitive science. The present work aims to shed light on cerebellar contributions to sensory feedback prediction mechanisms in auditory processing. The cerebellum is well-known for its role in the anticipation of somatosensory feedback following movements, but how far this function can be generalised across modalities, and how much it is tied to self-generated movements as opposed to more general sensory consequences, is less well understood. Understanding the role of the cerebellum during predictive auditory processing can serve as a guide for how to approach the more general questions of cerebellar contributions to cognition, the evolution of these capacities and the consequences of their potential disruption. It can also help us to form new predictions about the relationship between sensorimotor processing and cognition. To be able to anticipate the auditory consequences of motor acts is especially interesting in the context of vocal learning, in which case the function of sensory feedback prediction goes beyond being predicting a mere side-effect of a movement, but instead becomes crucial for the evaluation of the accuracy of the vocalisation (Torgeir Moberget & Ivry, 2016). Whether or not self-generated speech in this case employs different mechanisms from externally generated speech is an open question which we aim to explore further.

Research on the cerebellum has recently gained a lot of momentum, as its role as an important player during a wide variety of cognitive processes has become more established. A recent comprehensive investigation on cerebellar activation during a wide variety of motoric, cognitive and affective test conditions has highlighted the involvement of the cerebellum in a diverse set of processes such as language comprehension, autobiographical recall, and mental arithmetic (King, Hernandez-Castillo, Poldrack, Ivry, & Diedrichsen, 2019).

The cerebellum has furthermore been in the spotlight of evolutionary research concerning recent developments in the evolution of modern human cognition, which makes it an interesting candidate of study for understanding the evolutionary history of modern human specific capacities such as language (Barton & Venditti, 2014; Gunz et al., 2019).

The well-established consensus about cerebellar function during motor coordination suggests that the cerebellum generates internal forward models. These are thought to serve as predictors of anticipated sensory feedback of movements in order to compensate for the inherent timing delay that comes with sensory feedback (Albus, 1971; Marr, 1969).

The general idea that this concept could be applied not only to cerebellar motor function but also to the role the cerebellum plays in cognition have been articulated on many occasions (famously by Ito, 2008), but the specific dynamics of such a mechanism are yet to be understood clearly. To study the application of internal forward models supported by the cerebellum and their role in non-motor functions, we can turn to auditory language processing as an example of cerebellar involvement in sensory processing in the absence of movement.

We will present a systematic review of the state of the art of our knowledge on the role of the cerebellum during auditory processing. We will pay special attention to insights that can be gained from clinical research on auditory hallucinations. It is a widely held belief that the sense of agency of one's own actions depends on accurate prediction of the anticipated sensory feedback to those actions, and that this function relies in part on the (Moberget et al., 2018). The malfunctioning of these predictive processes may further be linked to a specific type of auditory hallucination involving disrupted sensory feedback prediction (Andreasen & Pierson, 2008). We evaluate the current evidence on a biological, mechanistic/computational and behavioural level in order to gain a clearer understanding of how the medical literature can inform our theories of cerebellar involvement in auditory processing and sensory processing generally, and we highlight the different levels on which this affects language behaviour and cognition more generally.

Acknowledgements

This research was supported by the FI 2019 grant (3 year grant for completion of PhD project) from the Generalitat de Catalunya, Spain.

References

Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*(1), 25–61. https://doi.org/10.1016/0025-5564(71)90051-4

Andreasen, N. C., & Pierson, R. (2008). The Role of the Cerebellum in Schizophrenia. *Biological Psychiatry*, *64*(2), 81–88. https://doi.org/10.1016/j.biopsych.2008.01.003

Barton, R. A., & Venditti, C. (2014). Rapid evolution of the cerebellum in humans and other great apes. *Current Biology: CB*, *24*(20), 2440–2444. https://doi.org/10.1016/j.cub.2014.08.056

Gunz, P., Tilot, A. K., Wittfeld, K., Teumer, A., Shapland, C. Y., van Erp, T. G. M., … Fisher, S. E. (2019). Neandertal Introgression Sheds Light on Modern Human Endocranial Globularity. *Current Biology*, *29*(1), 120-127.e5. https://doi.org/10.1016/j.cub.2018.10.065

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews. Neuroscience*, *9*(4), 304–313. https://doi.org/10.1038/nrn2332

King, M., Hernandez-Castillo, C. R., Poldrack, R. A., Ivry, R. B., & Diedrichsen, J. (2019). Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nature Neuroscience*, *22*(8), 1371–1378. https://doi.org/10.1038/s41593-019-0436-x

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*(2), 437-470.1.

Moberget, T., Doan, N. T., Alnæs, D., Kaufmann, T., Córdova-Palomera, A., Lagerberg, T. V., … Westlye, L. T. (2018). Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: A multisite mega-analysis of 983 patients and 1349 healthy controls. *Molecular Psychiatry*, *23*(6), 1512–1520. https://doi.org/10.1038/mp.2017.106

Moberget, T., & Ivry, R. B. (2016). Cerebellar contributions to motor control and language comprehension: Searching for common computational principles. *Annals of the New York Academy of Sciences*, *1369*(1), 154–171. https://doi.org/10.1111/nyas.13094

APPEALS TO ‘THEORY OF MIND’ NO LONGER EXPLAIN MUCH IN LANGUAGE EVOLUTION

JUSTIN SULIK*[1] and GARY LUPYAN[2]

*Corresponding Author: justin.sulik@gmail.com
[1]Cognition, Values and Behaviour, Ludwig Maximilian University of Munich, Germany
[2]Department of Psychology, University of Wisconsin–Madison

Theory of Mind (ToM) — the ability to represent or reason about others’ mental states (Apperly, 2012) — carries a significant burden in explaining how humans communicate (Woensdregt & Smith, 2017; Tomasello, Carpenter, Call, Behne, & Moll, 2005; Dunbar, 2004), especially when it comes to inferring someone’s communicative intentions in the absence of a conventional language. If we accept that language use requires sophisticated pragmatics (Scott-Phillips, 2014), then ToM must be either a necessary condition for language evolution or something that co-evolved with language (Woensdregt & Smith, 2017; Heyes, 2018; Sterelny, 2012; Levinson, 2006). In either case, it is an important component of the study of language evolution.

However, there is increasing recognition that ‘Theory of Mind’ does not refer to a single cognitive mechanism, and that ToM research employs a variety of methods that tap numerous distinct cognitive mechanisms (Apperly, 2012; Schaafsma, Pfaff, Spunt, & Adolphs, 2015). Recent empirical work demonstrates low levels of agreement between various individual-differences scales that are meant to tap ToM ability (Warnell & Redcay, 2019; Gernsbacher & Yergeau, 2019). Thus, the role of ToM in human communication — and by implication, in language evolution — is radically underspecified (Irvine, 2018).

We report two studies that aim to improve this state of affairs. Study 1 examined relationships between various measures of ToM ability. Study 2 examined whether ToM measures and several other non-ToM-related problem solving abilities predicted performance on a linguistic signaling task.

In Study 1, we administered a battery of ToM individual-differences measures to 50 adults. We included a False Belief measure (Valle, Massaro, Castelli, & Marchetti, 2015) — the gold-standard test of ToM ability (Apperly, 2012); a co-ordination task (Mehta, Starmer, & Sugden, 1994); a Keynesian beauty contest (Keynes, 1938); the ‘Understanding Others’ subscale of the Autistic Spectrum Quotient (Stewart & Austin, 2009; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001); and a task based on the Rational Speech Act (RSA) framework, in which people produce and interpret visual cues to disambiguate between a set of referents (Goodman & Frank, 2016). We found low, non-significant correlations between most of the measures, even though they all ostensibly involve reasoning about others' minds. This was also true of the False Belief task, despite its purported centrality to ToM. One exception was a Schelling coordination task (Mehta et al., 1994), success on which required participants to leverage salient features of shared knowledge. This task had moderate correlations with several ToM measures (even though the ToM measures did not correlate with each other), though notably not with the False Belief task.

In study 2 (n=106), we examined whether various ToM measures predicted performance on a linguistic signaling task which required taking the perspective of the one's interlocutor (Sulik & Lupyan, 2018). Because a pilot study showed that certain problem-solving skills (creative: Bowden & Jung-Beeman, 2007; analytic: Frederick, 2005) positively correlated with performance on this signaling task, we included assessments of these two reasoning measures in this study. This also allowed us to examine whether analytic and creative problem-solving explained any of the covariance between ToM measures. Although there were significant zero-order correlations between performance on the perspective-taking task and the ToM measures, the first-order partial correlations (controlling for the effect of analytic problem solving) were not significant, except for the first-order correlation between RSA task and Schelling task. For the remaining tasks, the apparent relationship is largely driven by analytic problem solving ability. In addition, creative problem solving predicted unique variance in the perspective-taking task, but not the other ToM tasks. Thus, most common tasks in this field lack an element crucial for the linguistic perspective-taking task — creativity — and this represents an important gap in the current literature on the evolution of pragmatics.

The ability to take the perspective of others — frequently encompassed under the umbrella term 'Theory of Mind' — is claimed to be central to the human ability to learn and use language (Levinson, 2006). Here, we found that measures purporting to assess ToM did not correlate with one another (Study 1) — a surprising finding if ToM tasks tap into a single general ToM trait. One exception was that the ability to discover salient coordination points (as assessed by the Schelling task), was correlated with several ToM measures, hinting that something like "salience reasoning" may be a better description of the relevant cognitive mechanism. Further, performance on ToM tasks — while varying between individuals — did not predict performance on a linguistic perspective-taking task (Study 2). Instead, better perspective taking was predicted by better performance on non-ToM tasks such as creative and analytic reasoning. Taken together, our results challenge the explanatory power accorded to ToM in human communication. We argue that to explain the evolution of human pragmatic inference, we need a better understanding of relevant cognitive mechanisms. We have identified several candidates, including creative reasoning and analytic reasoning.

References

Apperly, I. A. (2012). What is “theory of mind”? concepts, cognitive processes and individual differences. *The Quarterly Journal of Experimental Psychology*, *65*(5), 825–839.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (aq): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of autism and developmental disorders*, *31*(1), 5-17.

Bowden, E. M., & Jung-Beeman, M. (2007). Methods for investigating the neural components of insight. *Methods*, *42*, 87-99.

Dunbar, R. I. (2004). Gossip in evolutionary perspective. *Review of general psychology*, *8*(2), 100.

Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, *19*(4), 25-42.

Gernsbacher, M. A., & Yergeau, M. (2019). Empirical failures of the claim that autistic people lack a theory of mind. *Archives of Scientific Psychology*, *7*, 102–118.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818–829.

Heyes, C. (2018). Enquire within: cultural evolution and cognitive science. *Phil. Trans. R. Soc. B*, *373*(1743), 20170051.

Irvine, E. (2018). When is a code not a code? In C. Cuskley, M. Flaherty, H. Little, L. McCrohon, A. Ravignani, & T. Verhoef (Eds.), *The evolution of language: Proceedings of the 12th international conference (evolangxii).* NCU Press.

Keynes, J. M. (1938). *The general theory of employment, interest, and money.* New York: Harcourt Brace and Co.

Levinson, S. C. (2006). On the human “interaction engine”. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of human sociality: Culture, cognition and interaction* (pp. 39–69). Oxford: Berg.

Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *The American Economic Review*, *84*(3), 658-673.

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in cognitive sciences*, *19*(2), 65–72.

Scott-Phillips, T. C. (2014). *Speaking our minds.* London: Palgrave Macmillan.

Sterelny, K. (2012). Language, gesture, skill: The co-evolutionary foundations of language. *Phil. Trans. R. Soc. B*, *367*(1599), 2141–2151.

Stewart, M. E., & Austin, E. J. (2009). The structure of the autism-spectrum quotient (aq): Evidence from a student sample in scotland. *Personality and*

Individual Differences, *47*, 224-228.

Sulik, J., & Lupyan, G. (2018). Perspective taking in a novel signaling task: effects of world knowledge and contextual constraint. *Journal of Experimental Psychology: General*, *147*(11).

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences*, *28*(5), 675–691.

Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: the growing complexity of recursive thinking ability. *Europe's Journal of Psychology*, *11*(1), 112.

Warnell, K. R., & Redcay, E. (2019). Minimal coherence among varied theory of mind measures in childhood and adulthood. *Cognition*, *191*, 103997.

Woensdregt, M., & Smith, K. (2017). Pragmatics and language evolution. In *Oxford research encyclopedia of linguistics.*

REDUNDANT MORPHOLOGICAL MARKING BENEFITS CHILD LEARNERS

SHIRA TAL[*1] and INBAL ARNON[1]

[*]Corresponding Author: shira.tal1@mail.huji.ac.il
[1]Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel

Redundancy in morphological marking is found across languages (e.g., agreement (Haig & Forker, 2018). Its presence is somewhat puzzling given that it can add complexity to the language (Lupyan & Dale, 2010) and is dispreferred by speakers in production (Frank & Jaeger, 2008; Kurumada & Jaeger, 2015). What could be the functionality of redundant morphological marking? We propose learning as a possible explanation: redundant cues may facilitate learning, making them advantageous in the system as a whole. In line with this, the presence of multiple cues has been shown to facilitate learning across domains (Sloutsky & Robinson, 2013; Yoshida & Smith, 2005). Here, we ask whether similar facilitation occurs for multiple *morphological* cues, when the cues themselves have to be learned, specifically, when combining case-marking and word order to mark thematic assignment. If redundant morphological marking is facilitative, we should see improved learning despite the added complexity of learning an additional cue. Supporting this, cross-linguistic studies show that children's comprehension benefits from redundant morphological cues to thematic assignment (Chan, Lieven, & Tomasello, 2009; Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008). However, in many of these cases, the redundant form is also the prototypical and most frequent form in child-directed speech (Ibbotson & Tomasello, 2009), making it unclear whether comprehension was facilitated because of the redundant cues, or because of the greater frequency of these structures.

Here, we use an artificial language learning paradigm to compare the learnability of two linguistic systems: with and without a redundant case-marking morpheme. 60 Hebrew-speaking children (mean age 7;10) were exposed to one of two versions of the language (N=30 in each condition): one in which fixed word order alone serves as a cue for thematic assignment (non-Hebrew like OSV) and one in which both fixed word order and object case-marking serve as cues (only the object was case-marked). Following exposure, we asked children to match pictures to novel sentences (comprehension), and describe previously unseen pictures (production). If redundant marking helps learning, then children should show better learning in the redundant-condition, despite its greater complexity. Children successfully comprehended the language (better than chance in both conditions, $p<.0001$). As predicted, children showed better learning in the redundant-condition (91% vs. 65%, $p<.0001$, Figure 1). Importantly, children in this condition did notice the case-marking cue: when asked to choose between sentences with and without case-marking, they preferred those with case-marking (88% of the time). They also used case-marking in the majority of their productions (85% of the time). Case-marking also facilitated production: despite having to use an additional element, word order was more accurate when case-marking was produced ($p=.0001$). In an additional study (N=30, mean-age 7;10) we show that comprehension was still improved, but less so (relative to the control) when case-marking appeared only during test, indicating that redundancy during exposure was facilitative.

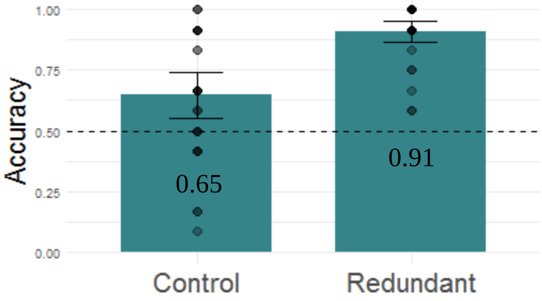


Figure 1. Accuracy scores by language condition. The dashed line indicates the chance level; error bars indicate confidence intervals; individual points indicate by-participant means.

Taken together, these findings suggest that redundant morphological cues can be facilitative for children. We are currently conducting follow-up work to ask whether redundancy impacts children and adults differently, as predicted by the *linguistic niche hypothesis* (Lupyan & Dale, 2010). Overall, the results suggest that redundancy can be facilitative in learning situations (at least for certain learners), and provide initial support for the idea that learning constraints help maintain redundancy in language.

Acknowledgements

We wish to thank Eitan Grossman for his helpful comments, as well as Kenny Smith and Jennifer Culbertson for helpful discussions and for their generous providing of stimuli. We also want to thank the research assistants at the Living Lab in the Bloomfield Science Museum (Rana Abu-Zhaya, Tali Cohen, Niva Goldberg), the museum staff, and the children and parents who participated in the studies. The first author acknowledges the scholarship from the Mandel Scholion Interdisciplinary Research Center in the Humanities and Jewish Studies at the Hebrew University of Jerusalem.

References

Chan, A., Lieven, E., & Tomasello, M. (2009). Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics*, *20*(2), 267–300. https://doi.org/10.1515/COGL.2009.015

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German Children's Comprehension of Word Order and Case Marking in Causative Sentences. *Child Development*, *79*(4), 1152–1167.

Frank, A. F., & Jaeger, T. F. (2008). Speaking Rationally : Uniform Information Density as an Optimal Strategy for Language Production. *The 30th Annual Meeting of the Cognitive Science Society (CogSci08)*, 939--944.

Haig, G., & Forker, D. (2018). Agreement in grammar and discourse : A research overview. *Linguistic Typology*, *56*(4), 715–734.

Ibbotson, P., & Tomasello, M. (2009). Prototype constructions in early language acquisition. *Language and Cognition*, *1*, 59–85. https://doi.org/10.1515/LANGCOG.2009.004

Kurumada, C., & Jaeger, T. F. (2015). Communicative efficiency in language production : Optional case-marking in Japanese. *Journal of Memory and Language*, *83*, 152–178. https://doi.org/10.1016/j.jml.2015.03.003

Lupyan, G., & Dale, R. (2010). Language Structure Is Partly Determined by Social Structure. *PLoS ONE*, *5*(1). https://doi.org/10.1371/journal.pone.0008559

Sloutsky, V. M., & Robinson, C. W. (2013). Redundancy Matters: Flexible Learning of Multiple Contingencies in Infants. *Cognition*, *126*(2), 156–164. https://doi.org/10.1016/j.cognition.2012.09.016.Redundancy

Yoshida, H., & Smith, L. B. (2005). Linguistic Cues Enhance the Learning of Perceptual Cues. *Psychological Science*, *16*(2), 90–95. https://doi.org/10.1111/j.0956-7976.2005.00787.x

COMPLEXITY CAN BE MAINTAINED IN SMALL POPULATIONS: A MODEL OF LEXICAL VARIABILITY IN EMERGING SIGN LANGUAGES

BILL THOMPSON*[1], LIMOR RAVIV[2], and SIMON KIRBY[3]

*Corresponding Author: billdthompson@berkeley.edu
[1]Social Science Matrix, University of California, Berkeley, California, USA
[2]Psychology of Language Dept., Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
[3]Centre for Language Evolution, The University of Edinburgh, Scotland, UK

Research on emerging sign languages suggests that social structure affects the process of language convergence (Meir, Israel, Sandler, Padden, & Aronoff, 2012). Specifically, sign languages that emerge in small, highly-connected communities are less conventionalized, showing greater lexical variability between speakers. On the other hand, languages that emerge in larger and sparser communities tend to be more uniform. This finding is somewhat surprising in light of theoretical results suggesting that shared conventions emerge *faster* in smaller populations (Baronchelli, Felici, Loreto, Caglioti, & Steels, 2006).

In this paper, we argue that the evidence from emerging sign language can be explained by an interaction between population size and our capacity to remember individual speakers/signers. Put simply, in the early stages of language formation, community members may employ two potential strategies in order to successfully interact with each other: memorize each others unique lexical variants, or try to align on a shared language. Importantly, the efficacy of these strategies and the ease with which they can be employed will vary in different population contexts.

Our hypothesis is that members of small communities are better able to keep track of each others variants, allowing them to successfully communicate with each other without the need to converge on a single variant at all. In contrast, such a strategy is much harder to maintain in larger groups with many more individuals variants to keep track of. Members of larger communities are therefore under a stronger pressure to reduce variability and converge on a shared lexical form. We hypothesize that when memory constraints are taken into account, rather than lexical convergence proceeding more rapidly in small populations, there will be situations where small populations preserve high levels of variability for longer – explaining the data we see in emerging sign languages.

We tested this hypothesis by simulating interaction in populations of language learners. In our model, each individual remembers the lexical variants used by specific individuals they have encountered, but also represents lexical variation in the population as a whole. We analysed the process of conventionalisation on a shared lexical form under the assumption that individuals combine these sources of information using hierarchical Bayesian inference. Under this model, learners draw on individual-specific representations when interacting with somebody familiar, but draw on a population-level generalisation when interacting with a stranger. We made the simplifying assumption that lexical variants can be represented in a one-dimensional continuous space, and that the distributions maintained by individuals can be approximated by Gaussian distributions. We measured conventionalisation (i.e. loss of lexical variation) as the variance in the lexical forms in the population.

Our analysis shows that, in important parts of the parameter space, memory limitations lead to an inverse relationship between population size and lexical variance. Small populations end up with languages that are highly variable, while larger populations converge on a uniform language. In contrast, a simpler variant of the model where agents do not keep track of who they are interacting with recapitulates the opposite relationship between population size and convergence suggested by (Baronchelli et al., 2006). Our results support the idea that simpler (i.e., more compressible) communication systems evolve in the presence of information-processing bottlenecks (Kirby, Tamariz, Cornish, & Smith, 2015), and are in line with the hypothesis that convergence in the early stages of language formation is driven by group size (Meir et al., 2012). Our findings also resonate with the idea that interacting with more strangers is an important factor in driving languages to be systematic and predictable (Wray & Grace, 2007).

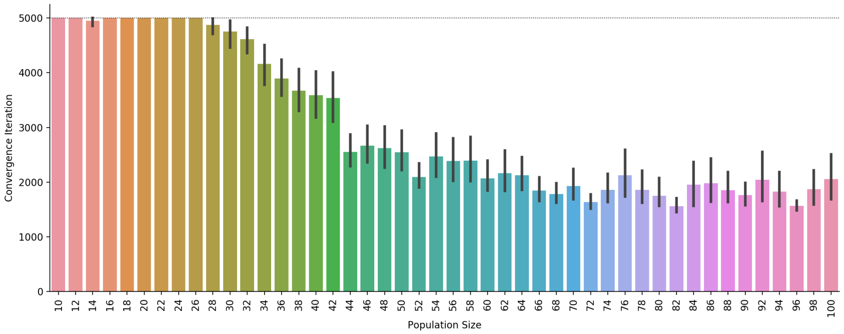


Figure 1. Time to convergence (y-axis) on a shared lexicon as a function of population size (x-axis) in simulated communities. Simulations ran for 5000 iterations: bars exceeding 5000 did not converge by the end of the simulation.

References

Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., & Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics: Theory and Experiment*, *2006*(06), P06014.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87 - 102.

Meir, I., Israel, A., Sandler, W., Padden, C. A., & Aronoff, M. (2012). The influence of community on language structure: Evidence from two young sign languages [Journal Article]. *Linguistic Variation*, *12*(2), 247-291.

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers: Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543–578.

CO-EVOLUTION OF CULTURE AND MEANING REVEALED THROUGH LARGE-SCALE SEMANTIC ALIGNMENT

BILL THOMPSON*[1], SEÁN G. ROBERTS[2], and GARY LUPYAN[3]

*Corresponding Author: billdthompson@berkeley.edu
[1]Social Science Matrix, University of California, Berkeley, USA
[2]ENCAP, Cardiff University, UK; Anthropology and Archaeology, University of Bristol, UK
[3]Department of Psychology, University of Wisconsin, Madison, USA

Do natural languages evolve to reflect the objective structure of the world (Gleitman & Fisher, 2005; Snedeker & Gleitman, 2004) or do they impose their own structure, with each language adapting to local communicative needs (Evans & Levinson, 2009; Davidson, 1973)? If languages mirror the objective structure of the world, words referring to natural kinds, common artifacts, and universal human actions should mean the same things in different languages. In contrast, if different languages impose their own structure, carving joints *into* nature, word meanings may exhibit substantial variability between languages, making cross-linguistic semantic alignment more difficult. This would not be surprising for specialised artefacts, regional animals or distinctions that relate to specific local conditions (e.g. distinctions between "ice" and "snow" are more likely in colder climates, Regier, Carstensen, & Kemp, 2016). However, the extent of the alignment between more common meanings (common animals and artifacts, natural features, quantifiers, body parts, and common verbs) is an open question. For example, do the English words 'five', 'near', and 'arm' mean the same thing as the Spanish words 'cinco', 'cerca" and 'brazo', respectively?

Quantifying semantic structure is difficult because word meanings are not directly observable (Cuyckens, Dirven, & Taylor, 2009). Here, we present a large-scale analysis of word meanings by taking advantage of recent advances in distributional semantics using machine-learning on natural language text. We obtained translation equivalents for 1,016 concepts in 74 languages using the NorthEuraLex dataset (Dellert & Jäger, 2017). We began by deriving within-language word-to-word similarities using the fast-text skipgram algorithm trained on language-specific versions of Wikipedia (Bojanowski, Grave, Joulin, & Mikolov, 2017). We also replicated on word embeddings derived from the OpenSubtitles database (Lison & Tiedemann, 2016) and a combination of Wikipedia and the Common Crawl dataset (Grave, Bojanowski, Gupta, Joulin, & Mikolov, 2018)). To compute semantic alignments for meaning c for language pair L_i and L_j, we first found the closest k semantic neighbours of c of L_i along with their context similarity score.

For example, the closest neighbours to the English word ‘beautiful’ are ‘colourful’ (.55), ‘love’ (.53) and ‘delicate’ (.51). We then found the translations of these neighbours in L_j and their corresponding proximity to the translation of c. The directional semantic alignment $L_i \to L_j$ is the correlation between c’s similarity to these neighbours in both languages. For example, the French translations of these neighbours are more distant from ‘beau’ (‘multicolore’=.22, ‘aimer’=.32 and ‘fin‘=.2), while other words are closer (‘frère’, ‘père‘) so the alignment is low. This was repeated in the opposite direction: the k closest semantic neighbours to c in L_j were identified and matched to their translations in L_i; the same correlation was calculated for $L_i \to L_j$. Final semantic alignment is the average of these two correlations. We validated the measure by correlating it with human translatabiliy judgements (e.g., Tokowicz et al., 2002; Allen & Conklin, 2014).

The most alignable meanings across languages stand out not as being especially concrete or reflecting natural joints, but as domains that have high internal coherence such as number words and kinship terms. In comparison, words for common artifacts, actions, and natural kinds have much lower alignments indicating that these words have different semantic neighborhoods in different languages.

If languages reflect cultural factors, then languages should be more aligned if they are spoken by people with similar cultures. We confirmed that cultural similarity (the proportion of cultural traits in common based on 92 non-linguistic cultural traits for 39 societies, Kirby et al., 2016) predicted semantic alignment between languages, even when controlling for historical relatedness and geographic proximity (b= 0.2, $\chi^2(1)$= 16.56, $p <$.001). Cultural similarity related to subsistence type was correlated with semantic alignment in domains including ‘food and drink’ (r = .3), ‘animals’ (r = .29), ‘agriculture and vegetation’ (r = .25), ‘clothing and grooming’ (r = .25), ‘social and political relations’ (r = .15), and ‘spatial relations’ (r = .1, all adjusted p-values $<$ .05). These reflect well-known relations between subsistence types and culture (Murdock & Provost, 1973; Sellen & Smay, 2001; Peoples & Marlowe, 2012; Botero et al., 2014; Gavin et al., 2018; Majid et al., 2018). This indicates that cultural and historical processes influence the evolution of natural language semantics. Consistent with the idea that languages that emerge in larger communities have more systematic structure (Raviv, Meyer, & Lev-Ari, 2019), we find that semantic alignment positively associated with population size. Controlling for shared history, languages spoken by larger groups tend to align better with one another (b = .002, t = 7.1, $p <$ .001).

Our results show that even frequent concrete meanings show substantial cross-linguistic differences – differences which are predictable from shared culture and history. Despite some of the shortcomings of corpus-derived semantics (which makes our analysis more conservative), we believe the present work provides a major step forward for understanding the evolutionary factors that shape the emergence and evolution of linguistic meaning, and particularly the impact of shared culture (Thompson et al., 2016).

References

Allen, D., & Conklin, K. (2014). Cross-linguistic similarity norms for Japanese–English translation equivalents. *Behavior Research Methods*, *46*(2), 540-563.

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146.

Botero, C. A., Gardner, B., Kirby, K. R., Bulbulia, J., Gavin, M. C., & Gray, R. D. (2014). The ecology of religious beliefs. *Proceedings of the National Academy of Sciences*, *111*(47), 16784–16789.

Cuyckens, H., Dirven, R., & Taylor, J. R. (2009). *Cognitive approaches to lexical semantics* (Vol. 23). Walter de Gruyter.

Davidson, D. (1973). On the Very Idea of a Conceptual Scheme. *Proceedings and Addresses of the American Philosophical Association*, *47*, 5-20. (ArticleType: research-article / Full publication date: 1973 - 1974 / Copyright © 1973 American Philosophical Association)

Dellert, J., & Jäger, G. (2017). *NorthEuraLex (version 0.9).*

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and brain sciences*, *32*(5), 429–448.

Gavin, M. C., Kavanagh, P. H., Haynie, H. J., Bowern, C., Ember, C. R., Gray, R. D., Jordan, F. M., Kirby, K. R., Kushnick, G., Low, B. S., Vilela, B., & Botero, C. A. (2018). The global geography of human subsistence. *Royal Society Open Science*, *5*(9), 171897.

Gleitman, L., & Fisher, C. (2005). Universal aspects of word learning. In J. McGilvray (Ed.), *The Cambridge Companion to Chomsky* (p. 123-142). New York, NY: Cambridge University Press.

Grave, E., Bojanowski, P., Gupta, P., Joulin, A., & Mikolov, T. (2018). Learning word vectors for 157 languages. In *Proceedings of the international conference on language resources and evaluation (lrec 2018).*

Kirby, K. R., Gray, R. D., Greenhill, S. J., Jordan, F. M., Gomes-Ng, S., Bibiko, H.-J., Blasi, D. E., Botero, C. A., Bowern, C., Ember, C. R., et al.. (2016). D-place: A global database of cultural, linguistic and environmental diversity. *PloS one*, *11*(7), e0158391.

Lison, P., & Tiedemann, J. (2016). Opensubtitles2016: Extracting large parallel corpora from movie and TV subtitles. In *Proceedings of the international conference on language resources and evaluation (lrec 2016).*

Majid, A., Roberts, S. G., Cilissen, L., Emmorey, K., Nicodemus, B., OGrady, L., Woll, B., LeLan, B., De Sousa, H., Cansler, B. L., Shayan, S., Vos, C. de, Senft, G., Enfield, N. J., Razak, R. A., Fedden, S., Tufvesson, S., Dingemanse, M., Ozturk, O., Brown, P., Hill, C., Guen, O. L., Hirtzel, V., Gijn, R. van, Sicoli, M. A., , & Levinson, S. C. (2018). Differential coding of perception in the worlds languages. *Proceedings of the National Academy of Sciences*, *115*(45), 11369–11376.

Murdock, G. P., & Provost, C. (1973). Factors in the division of labor by sex: A cross-cultural analysis. *Ethnology*, *12*(2), 203–225.

Peoples, H. C., & Marlowe, F. W. (2012). Subsistence and the evolution of religion. *Human Nature*, *23*(3), 253–269.

Raviv, L., Meyer, A., & Lev-Ari, S. (2019). Larger communities create more systematic languages. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1907), 20191262.

Regier, T., Carstensen, A., & Kemp, C. (2016). Languages support efficient communication about the environment: Words for snow revisited. *PloS one*, *11*(4), e0151138.

Sellen, D. W., & Smay, D. B. (2001). Relationships between subsistence and age at weaning in ”preindustrial” societies. *Human Nature*, *12*, 47–87. (Accessed on: 2019-01-18)

Snedeker, J., & Gleitman, L. (2004). Why is it hard to label our concepts? In D. G. Hall & S. Waxman (Eds.), *Weaving a Lexicon* (illustrated edition ed., p. 257-294). Cambridge, MA.: The MIT Press.

Thompson, B., Kirby, S., & Smith, K. (2016). Culture shapes the evolution of cognition. *Proceedings of the National Academy of Sciences*, *113*(16), 4530–4535.

Tokowicz, N., Kroll, J. F., De Groot, A. M., & Van Hell, J. G. (2002). Number-of-translation norms for dutchenglish translation pairs: A new tool for examining language production. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 435–451.

SEXUAL DIMORPHIC CHORUSING IN THE WILD INDRIS

VALERIA TORTI[1], DARIA VALENTE[1], CHIARA DE GREGORIO[1], ANNA ZANOLI[1], LONGONDRAZA MIARETSOA[1], CRISTINA GIACOMA[1], MARCO GAMBA[1].

*Corresponding Author: valeria.torti@unito.it
[1]Life Sciences and Systems Biology, University of Torino, Torino, Italy

Animals can produce vocal rhythms in an interactive, coordinated manner (Couzin 2018). Comparing structural, spectral and temporal features across species (Fitch 2000) may help in reconstructing the evolutionary history of human speech (Ravignani and Norton 2017). Singing primates (Geissmann, 2000), which produce elaborated and complex sequences of vocalizations, are of particular interest for this topic. Similarly to humans, indris (*Indri indri*) assemble simple units into more complex structures to convey different information. Individuals react differently to different songs, confirming the presence of functionally referential communication systems (Clark et al. 2006).

Indris are the only singing lemurs and emit songs whose most distinctive portions are “descending phrases”, made of 2-5 units. Mated indris have been reported to sing in pairs, to enhance pair-bonding and defend their territories (Pollock, 1986; Torti et al., 2013; Bonadonna et al., 2017). Songs may have the form of a chorus whenever the subadult members of the groups also utter their contribution in a precise and coordinated manner (Torti et al., 2018). Indri songs exhibit turn-taking between individuals of different sexes and a variable degree of overlap between group members (Gamba et al., 2016). Songs have various functions depending on the context in which are emitted (Torti et al., 2013), and they are used for both inter and intra-group communication.

We recorded spontaneous vocalizations of 8 groups of indris at the Maromizaha New Protected Area (18°56′S, 48°27′E), from 2008 to 2018. Focal animal sampling (Altmann, 1974) allowed the attribution of each vocal profile to a signaler. To investigate the timing and rhythm of songs, we measured the amount of co-singing between different individual contributions (percentage of overlap;

Gamba et al. 2014) and the inter onset intervals (IOIs, duration between the starting points of two successive notes in the same song; Gamba et al. 2016) of adjacent units (Sasahara et al., 2015). We extracted the pitch contour and labeled each unit using the phrase in which it was emitted and the sex of the emitter. We then calculated the similarity across different individual songs using the Levenshtein distance. Finally, we classified song units in phrases through DTW and clustering analyses (Gamba et al. 2018).

Our results show that: a) indris can synchronize their utterances showing non-random overlap between singers, with an overlapping rate of the pair contributions that changes according to the number of singers in the chorus; b) both dominant and non-dominant indris can coordinate their calls and there is evidence for an ability of precise timing during song emission; c) indri songs show the presence of sex dimorphism, both in the overall timing and repertoire size than in the unit and phrase structure, with females being more flexible than males in their contributions, and d) the structure of phrases possess individually distinctive characteristics.

In line with previous findings (De Gregorio et al., 2018), we observed that indris within a group coordinate on average more than 70% of their contributions suggesting that duetting is indeed associated with pair cohesion, as a proxy of the strength of the pair bond (Geissmann & Orgeldinger, 2000). The most consistent portion of the song, made of ascending or descending sequences of units, shows reliable timing and pitch variation, a crucial feature of birdsong and human speech (Levinson & Holler, 2014). We found support for our prediction that the phrase structure of songs varied between reproductive males and females. The presence of pitch sex dimorphism in nonhuman primate vocal signals is rare and is identified as a prerequisite in the evolution of human perceptual abilities (Patel, 2010). It appears that indri male's song has a more fixed pattern, whereas females could adjust their contribution, in agreement with findings on the white-cheeked gibbons N. leucogenys (Deputte, 1982), on baboons and Japanese macaques (Lemasson et al., 2011; Lemasson et al., 2016). Our results also show that strong individuality is encoded in the indris' phrases, thus the potential to provide conspecifics with emitter's identity cues.

Studies on rhythm and synchrony in primates have historically been focused on anthropoid species (Ravignani, 2019), with poor investigation on prosimian vocal behavior. Our study suggests that the indris are a good model for further investigations of the evolution of human speech features, because of the turn-taking between individuals and the variable degree of overlap, commonly identified as traits of modern human communication.

References

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour,* 49, 227–267.

Arnold K., & Zuberbühler K. (2006). Language evolution: semantic combinations in primate calls. *Nature*, 18, 441(7091):303.

Bonadonna G., Torti V., Sorrentino V., Randrianarison R.M., Zaccagno M., Gamba M., Tan C.L., Giacoma C. (2017). Territory exclusivity and intergroup encounters in *Indri indri* upon methodological tuning. *The European Zoological Journal*, 84(1), 238-251.

Couzin I.D. (2018). Synchronization: The Key to Effective Communication in Animal Collectives. *Trends in Cognitive Sciences,* 22(10), 844-846.

De Gregorio C., Zanoli A., Valente D., TortiV., Bonadonna G., Randrianarison R.M., Giacoma C., Gamba M. (2018). Female indris determine the rhythmic structure of the song and sustain a higher cost when the chorus size increases. *Current Zoology*, zoy058.

Deputte, B. (1982). Duetting in male and female songs of the white-checked gibbon (*Hylobates con-color leucogenys*). In C.T. Snowdon, C.H. Brown, & M. Peterses (Eds.), *Primate Communication* (pp. 67-93). Cambridge Univ. Press, Cambridge.

Fitch W.T. (2000). The Evolution of Speech: A Comparative Review. *Trends in Cognitive Sciences* , 4(7), 258–267.

Gamba M., Torti V., Estienne V., Randrianarison R.M., Valente D., Rovara P., Bonadonna G., Friard O., Giacoma C. (2016). The Indris Have Got Rhythm! Timing and Pitch Variation of a Primate Song Examined between Sexes and Age Classes. *Frontiers in Neuroscience*, 10: 249.

Gamba M., Torti V., Valente D., De Gregorio C., Friard O., Giacoma C. (2018). Primate songs and their relevance in the study of language evolution. In C. Cuskley, M. Flaherty, L. McCrohon, H. Little, A. Ravignani, T. Verhoef, *The Evolution of Language. Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)* (pp.134-136). Evolang 12 Organizing Committee.

Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In N. Wallin, B. Merker and S. Brown (Eds.), *The origins of music* (pp. 103-123). Cambridge, MA: MIT Press.

Geissmann T., & Orgeldinger, M. (2000). The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Animal Behaviour*, 60(6), 805-809.

Lemasson, A., Ouattara , K., Petit, E.J., & Zuberbühler, K., (2011). Social learning of vocal structure in a nonhuman primate? *BMC Evolutionary Biology*, 11:362.

Lemasson, A., Jubin, R., Masataka, N., & Arlet, M. (2016). Copying hierarchical leaders' voices? Acoustic plasticity in female *Japanese macaques*. *Scientific Reports*, 6:21289.

Levinson S.C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of The Royal Society B Biological Sciences*, 369(1651).

Marler, P. (2000). Origins of music and speech: Insights from animals. In N. Wallin, B. Merker and S. Brown (Eds.), *The Origins of Music* (pp. 31– 48). Cambridge, MA: MIT Press.

Patel, A.D. (2010). Music, biological evolution, and the brain. In M. Bailar (Eds.), Emerging Disciplines (pp. 91-144). Houston, TX: Rice UP.

Pollock, J.I. (1986). The song of the Indris (*Indri indri*; Primates: Lemuroidea): natural history, form and function. *International Journal of Primatology,* 7, 225-267.

Ravignani A., Norton P. (2017). Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution*, 2(1), 4-19.

Ravignani, A. (2019). Rhythm and synchrony in animal movement and communication. *Current Zoology*, 65(1), 77–81.

Sasahara, K., Tchernichovski, O., Takahasi, M., Suzuki, K., and Okanoya, K. (2015). A rhythm landscape approach to the developmental dynamics of birdsong. Journal of the Royal Society Interface, 12:20150802.

Torti V., Gamba M., Rabemananjara Z.H., Giacoma C. (2013). The song of the indris: contextual variation in the long distance calls of a lemur. *Italian Journal of Zoology*, 80(4), 596-607.

Torti V., Valente D., De Gregorio C., Comazzi C., Miaretsoa L., Ratsimbazafy J., Giacoma C., Gamba M. (2018). Call and be counted! Can we reliably estimate the number of callers in the indri's (Indri indri) song? *PLoS ONE,* 13(8).

CHUNKING AND ASSOCIATIVE LEARNING IN NON-HUMAN PRIMATES

LAURE TOSATTO*[1], JOEL FAGOT[1], DEZSO NEMETH[2,3,4], and ARNAUD REY[1]

*Corresponding Author: laure.tosatto@univ-amu.fr
[1]Laboratoire de Psychologie Cognitive, CNRS & Aix-Marseille Université, Marseille France
[2]CRNL – Lyon Neuroscience Research Center, Université Claude Bernard Lyon 1, Lyon, France
[3]Institute of Psychology, Eotvos Lorand University, Budapest, Hungary
[4]Hungarian Academy of Sciences, Budapest, Hungary

1. Introduction

A key process in language acquisition is the ability to extract sequences of units (phonemes, words, etc.) that occur together regularly and repeatedly in spoken or written language (Bannard & Matthews, 2008).

McCauley and Christiansen (2019) recently introduced a new model of language comprehension and in which chunking mechanisms are supposed to play a central role in the extraction of recurring multiword units.

Several theories and computational models suggest that chunking mechanisms are more generally central in sequence learning, as the units composing these sequences are in fact associated through elementary associative or Hebbian learning mechanisms and compiled into chunks of information (e.g., the self-organizing consciousness theory, Perruchet & Vinter, 2002; Parser, Perruchet & Vinter, 1998; TRACX, French, Addyman & Mareschal, 2011). Chunking and associative learning mechanisms are not uniquely human, suggesting that the fundamental mechanisms involved in language learning are shared by many other animal species.

Non-human animals have indeed been shown form chunks while learning sequences (e.g., Terrace, 1987), but we know less about how these chunks are formed and evolve during practice. Studying sequence learning behaviors in animals is therefore essential for assessing the similarities and differences between human and non-human animals and reaching a better understanding of chunking mechanisms and their role in language acquisition (Rey, Minier, Malassis, Bogaerts, & Fagot, 2019).

2. Method

Using an operant conditioning device (Fagot & Bonté, 2010), a total of eighteen Guinea baboons (*Papio papio*) were initially trained to produce random visuo-motor sequences by touching a moving target red circle on a touch screen.

They were then presented with a repeated sequence of nine positions and had to perform a serial response time task on the touch screen by touching the moving target. They produced this same motor sequence during 1000 successive trials.

3. Results

We interpreted decrease in response time between two successive positions as these positions being chunked together, whereas increase in response time was interpreted as the chunk boundary. Thus, we identified chunking patterns of the sequence for every baboon. Additionally, the evolution of response times revealed that these patterns evolved during the course of learning, from a concatenation of initially small chunks into larger chunks later on.

4. Discussion

These results provide new evidence on the dynamics of chunking processes in non-human primates and, more generally, on the mechanisms involved in implicit statistical learning, a core learning process in human language acquisition. They suggest that chunking mechanisms start small (chunks of 2 to 3 elements) as it was previously found in humans (e.g. Verwey, 2001; Wymbs, Bassett, Mucha, Porter & Grafton, 2012). We also found that these small chunks were later concatenated into larger chunks leading to increasingly compressed forms of information. These elementary mechanisms are certainly also present and crucial in building the core elements of language comprehension and production as suggested by McCauley and Christiansen (2019).

Acknowledgements

This research is funded by the "Chunked" project (ANR-17-CE28-0013) and supported by the Institute of Language Communication and the Brain (ILCB).

References

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19*(3), 241-248.

Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (Papio papio). *Behavior research methods*, *42*(2), 507-516.

French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based connectionist framework for sequence segmentation and chunk extraction. *Psychological Review*, *118*(4), 614.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in cognitive sciences*, *19*(3), 117-125.

McCauley, S. M., & Christiansen, M. H. (2019). Language learning as language use: A cross-linguistic model of child language development. *Psychological review*, *126*(1), 1.

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of memory and language*, *39*(2), 246-263.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences*, *25*(3), 297-330.

Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (2019). Regularity Extraction Across Species: Associative Learning Mechanisms Shared by Human and Non-Human Primates. *Topics in cognitive science*.

Terrace, H. S. (1987). Chunking by a pigeon in a serial learning task. *Nature*, *325*(6100), 149.

Verwey, W. B. (2001). Concatenating familiar movement sequences: The versatile cognitive processor. *Acta psychologica*, *106*(1-2), 69-95.

Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A., & Grafton, S. T. (2012). Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. *Neuron*, *74*(5), 936-946.

THE EVOLUTION OF SPATIAL DEVICES IN GESTURAL STORYTELLING

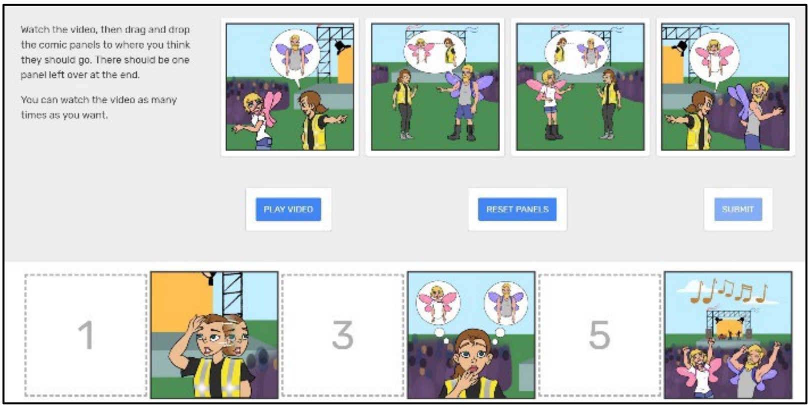
ROSS TOWNS*, MARCELLO A. GÓMEZ-MAUREIRA, KEES SOMMER, MAX VAN DUIJN AND TESSA VERHOEF

*Corresponding Author: r.j.towns@liacs.leidenuniv.nl
Creative Intelligence Lab, LIACS, Leiden University, Netherlands

A key communicative resource available to sign language users is the use of space to distinguish between referents and to express relationships between them. Signers can use space to convey a range of relational information, for example marking verb agreement by directing signs toward locations associated with distinct referents (Padden, 1988), or signaling shifts between 1st and non-1st person perspectives via shifts in bodily orientation (see Stec, 2013, for a review). While spatial tracking of referents is found in co-speech gesture (Perniss & Özyürek, 2015), evidence from young sign languages suggests that systematic use of spatial devices, or spatial modulation (Senghas & Coppola, 2001), emerges over successive cohorts of signers (Kocab, Pyers, & Senghas, 2014; Montemurro, Flaherty, Coppola, & Brentari, 2019) and is lacking in early stages of sign language emergence (Meir, Padden, Aronoff & Sandler, 2007).

We present an investigation of the cultural transmission of spatial devices using a novel experimental method. Drawing on artificial sign language experiments (Motamedi, Schouwstra, Smith, Culbertson, & Kirby, 2019) and the cultural evolution of stories (Bartlett, 1920; Mesoudi, Whiten, & Dunbar, 2006), we asked hearing non-signers in transmission chains to interpret and retell a short story narrated using improvised silent gesture.

We collected data from 54 participants in 9 chains of 6 generations each. The experiment was conducted over 3 days at a dedicated science area of a music festival. Participants first watched a video of the story depicted silently through pantomime before being presented with an incomplete 6 panel comic strip (see Fig. 1). To complete it, participants had to pick 3 out of 4 possible comic panels and arrange them according to their interpretation of the sequence of events depicted in the video. The story and reconstruction task were designed so that multiple orderings of panels were plausible. Participants were then asked to record their own pantomime videos, using their choice of panels as a prompt.



Watch the video, then drag and drop the comic panels to where you think they should go. There should be one panel left over at the end.

You can watch the video as many times as you want.

Figure 1. The interface of the story reconstruction task, operated via a touchscreen by dragging and dropping comic panels from the top right to the empty slots (marked 1, 3 and 5).

Participants at each generation viewed the video from the previous member of their chain, except at Generation 1, where participants watched a pre-recorded video in which an experimenter depicted a randomly selected story sequence in which all gestures were directed toward the camera and characters were identified using lexical labels (e.g. BEARD for the male character). The same seed video was used for all chains. During recording, participants saw a live video feed of themselves. In 5 out of 9 chains, both live and recorded video output were flipped horizontally to show a mirror image of the participant in order to identify a possible effect of visual feedback on participants' use of directional vs lexical labelling strategies.

Initial video coding finds that participants identified characters using a mixture of lexical labels and directional strategies, typically shifting their body orientation when embodying different characters. The use of this body shifting strategy increased over generations. Our preliminary findings are thus potentially in line with recent work by Motamedi, Schouwstra, Smith, Culbertson, & Kirby (2018), which found that participants in transmission chains used spatial gestures systematically to identify referents. However, in contrast to previous gesture transmission experiments, which found that without a pressure for efficiency imposed by dyadic communication, participants produced longer, more elaborate gesture sequences (Motamedi et al., 2019), we found that participants' gesture videos decreased in length over generations, despite the absence of dyadic communication in our experiment. This may be due to the visual presentation of our stimulus items providing a shared referential environment or common ground (Clark, 1996) across generations, allowing for efficient identification of characters from simple gestures. Further analysis will identify whether generational transmission in our chains led to systematic use of spatial gestures, and how efficiently this strategy was combined with lexical labelling.

References

Bartlett, A. F. C. (1920). Some Experiments on the Reproduction of Folk-Stories. *Folklore*, *31*(1), 30–47.

Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620539

Kocab, A., Pyers, J., & Senghas, A. (2014). Referential shift in Nicaraguan Sign Language: A transition from lexical to spatial devices. *Frontiers in Psychology*, *5*(OCT), 1–13. https://doi.org/10.3389/fpsyg.2014.01540

Meir, I., Padden, C. A., Aronoff, M., & Sandler, W. (2007). Body as subject. Journal of Linguistics, 43(3), 531-563.

Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, *97*(3), 405–431. https://doi.org/10.1348/000712605X85871

Montemurro, K., Flaherty, M., Coppola, M., & Brentari, D. (2019). Grammaticalization of the Body and Space in Nicaraguan Sign Language. In M. Brown & B. Dailey (Eds.), *Proceedings of the 43rd Boston University Conference on Language Development* (pp. 415–426). Cascadilla Press.

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J. & Kirby, S. (2018). The cultural evolution of spatial modulations in artificial sign languages. In Cuskley, C., Flaherty, M., Little, H., McCrohon, L., Ravignani, A. & Verhoef, T. (Eds.): The Evolution of Language: Proceedings of the 12th International Conference (EVOLANGXII). doi:10.12775/3991-1.078

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J., & Kirby, S. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition*, *192* (April), 103964. doi: 10.1016/j.cognition.2019.05.001

Padden, C. (1988) *Interaction of morphology and syntax in American Sign Language (Outstanding Dissertations in Linguistics, series IV)* New York: Garland Press

Perniss, P., & Özyürek, A. (2015). Visible cohesion: A comparison of reference tracking in sign, speech, and co-speech gesture. *Topics in Cognitive Science, 7*(1), 36–60. https://doi.org/10.1111/tops.12122

Senghas, A., & Coppola, M. (2001). Children Creating Language: How Nicaraguan Sign Language Acquired a Spatial Grammar. *Psychological Science*, *12*(4), 323–328. https://doi.org/10.1111/1467-9280.00359

Stec, K. (2013). Meaningful shifts. *Gesture*, *12*(3), 327–360. https://doi.org/10.1075/gest.12.3.03ste

INVESTIGATING THE CULTURAL EVOLUTION OF THE BLOMBOS AND DIEPKLOOF ENGRAVINGS

KRISTIAN TYLÈN*[1,2], RICCARDO FUSAROLI[1,2], SERGIO GONZALEZ DE LA HIGUERA ROJO[1], KATRIN HEIMANN[2], NICOLAS FAY[3], NIELS N. JOHANNSEN[2,4], FELIX RIEDE[2,4], and MARLIZE LOMBARD[5]

*Corresponding Author: kristian@cc.au.dk

[1] Department of Linguistics, Cognitive Science and Semiotics, Aarhus University, Denmark

[2] Interacting Minds Centre, Aarhus University, Denmark

[3] School of Psychological Science, University of Western Australia, Australia

[4] Department of Archaeology and Heritage Studies, Aarhus University, Denmark

[5] Centre for Anthropological Research, University of Johannesburg, South Africa

1. Introduction

The empirical study of language evolution either studies the sparse material remains of past expressive behaviors (archeology) or humans (genetics and cranial casts), or experimentally tests modern humans in simple laboratory experiment. A concern could be if our laboratory investigations are sufficiently connected to the phenomenon they are intended to investigate. Here we attempt to bridge this epistemic gap by using archeological items directly as stimuli in an experiment.

Dating back as far as 100 ka, the South African Blombos ochre and the Diepkloof ostrich egg engravings are considered among the earliest fossilized evidence of human symbolic behavior and have thus informed discussions about early language evolution (Henshilwood, d'Errico, & Watts, 2009; Hodgson, 2014; Texier et al., 2013). Of special interest is the temporal trajectory spanning more than 40 thousand years from earlier simpler parallel line patterns to later complex cross-hatchings (see figure 1A).

We hypothesize that this development is indicative of a cumulative adaptive evolution of engraving practices refining them over generations through processes of reproduction, transmission and learning to become increasingly adapted for their symbolic functions.

In order to investigate if the temporal development of the patterns indeed can be associated with mechanisms of cultural transmission, we first extracted a number of measures characterizing the compositional development of the original patterns, such as Kolmogorov and perimetric complexity (Kolmogorov, 1963; Tamariz & Kirby, 2015), Shannon entropy (Shannon, 1948), and symmetry (Kovesi, 1997). Generally, we observe that patterns become more symmetrical and more complex over time, however with a quadratic trend for complexity (intermediate period patterns are more complex). We then conducted an experimental simulation in the lab. Eight transmission chains of eight generations reproduced patterns in a digitized experimental environment. We seeded each chain with stylized outlines of engraved patterns from the earliest period of the Blombos or Diepkloof collections (see figure 1B). After a brief presentation (3 s) of a pattern the participant was instructed to reconstruct it from memory by placing and rotating lines using the computer mouse. The resulting figures were passed down as training stimulus for participants of the next generations etc. (Kirby & Hurford, 2002).

We then apply the same analytic metrics used to quantify the development of the original patterns to analyze the compositional development of the experimentally derived patterns. While data collections are still under completion, preliminary analyses suggest that indeed the developments of the experimentally derived patterns correlate with corresponding measures of the original patterns indicating that the Blombos and Diepkloof engravings could be the product of cumulative cultural transmission processes.

Figure 1. **A**: organization of the Blombos engravings as a function of time with the older items to the left dating to ~100 ka and the more recent to the right dating to ~60 ka. **B**: example of data from the iterated learning experiment. From left to right is the reproduction from generation 1-8.

Acknowledgements

The project has been supported by a seed grant (no21887) from the Interacting Minds Centre, Aarhus University and an International Network Program grant 5132-00026B from the Danish Agency for Science, Technology and Innovation.

References

Henshilwood, C. S., d'Errico, F., & Watts, I. (2009). Engraved ochres from the middle stone age levels at Blombos Cave, South Africa. *Journal of human evolution, 57*(1), 27-47.

Hodgson, D. (2014). Decoding the Blombos engravings, shell beads and Diepkloof ostrich eggshell patterns. *Cambridge Archaeological Journal, 24*(01), 57-69.

Kirby, S., & Hurford, J. R. (2002). The emergence of linguistic structure: An overview of the iterated learning model. In *Simulating the evolution of language* (pp. 121-147): Springer.

Kolmogorov, A. N. (1963). On tables of random numbers. *Sankhyā: The Indian Journal of Statistics, Series A, 25*(4), 369-376.

Kovesi, P. (1997). *Symmetry and asymmetry from local phase.* Paper presented at the Tenth Australian joint conference on artificial intelligence.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal, 27*(3), 379-423.

Tamariz, M., & Kirby, S. (2015). Culture: copying, compression, and conventionality. *Cogn Sci, 39*(1), 171-183.

Texier, P.-J., Porraz, G., Parkington, J., Rigaud, J.-P., Poggenpoel, C., & Tribolo, C. (2013). The context, form and significance of the MSA engraved ostrich eggshell collection from Diepkloof Rock Shelter, Western Cape, South Africa. *Journal of Archaeological Science, 40*(9), 3412-3431.

GESTURAL ORIGINS OF VERB AND NOUN ENCODING IN SIGN LANGUAGE EMERGENCE

TESSA VERHOEF*[1] and RYAN LEPIC[2]

*Corresponding Author: t.verhoef@liacs.leidenuniv.nl
[1]Institute of Advanced Computer Sciences, Leiden University, Netherlands
[2]Department of Psychology, University of Chicago, United States of America

Human languages use verbs and nouns distinguish actions vs. objects (Langacker 1987) and predication vs. reference (Croft 2000). Prototypically, verbs predicate about actions while nouns refer to objects. The visual modality affords the ability to iconically represent actions and objects with the movement and shape of the hands; in sign languages, related verbs and nouns can be distinguished by manner of movement (Supalla & Newport 1978, Johnston 2001, and Tkachman & Sandler 2013) or handshape (Padden et al. 2015). To examine the possible gestural origins of verb/noun-encoding in representations of actions/objects in sign language emergence, here we ask whether non-signers are sensitive to representational strategies that emulate verb/noun-encoding strategies in natural sign languages.

It has been found that sign-naïve gesturers show a strong preference for representing actions with "handling" handshapes, which show a human hand in action, and a slight preference for representing objects with "instrument" handshapes, which show human manipulation of the object, an encoding strategy that also distinguishes verbs and nouns in ASL (Verhoef et al. 2016). In ASL, movement patterns also distinguish some related verbs and nouns: verbs are formed with longer/continuous movements, while nouns are formed with faster/constrained movements (Supalla & Newport 1978). Here, we test the possible interplay between these handshape and movement preferences. We expect that non-signers will interpret gestures formed with handling handshapes and continuous movements as depicting actions, and gestures with instrument handshapes and constrained movements as referring to objects. However, it is not clear what non-signers may prefer when these mappings are in conflict (Table 1).

Table 1. Expected non-signer responses in four experimental conditions

	Handling handshape	*Instrument handshape*
Continuous movement	Favors **action** interpretation	Conflicting biases
Constrained movement	Conflicting biases	Favors **object** interpretation

We recruited 1175 participants via Crowdflower, and asked them to identify each of the gestures in a pair of videos as representing either an action (e.g. "using a handsaw") or object (e.g. "a handsaw"). Figure 1 shows the proportion of gestures labeled as referring to an action (as opposed to an object) for each gesture type, in the four conditions from Table 1. When movement is constant across gesture pairs and handshape varies, participants map handling handshapes to actions and instrument handshapes to objects. When handshape is constant and movement varies, participants map continuous movements to actions and constrained movements to objects. In the interaction conditions, participants weight their preferences for mappings based on handshapes over their preferences for mappings based on movement, suggesting that handshape is a more salient cue.

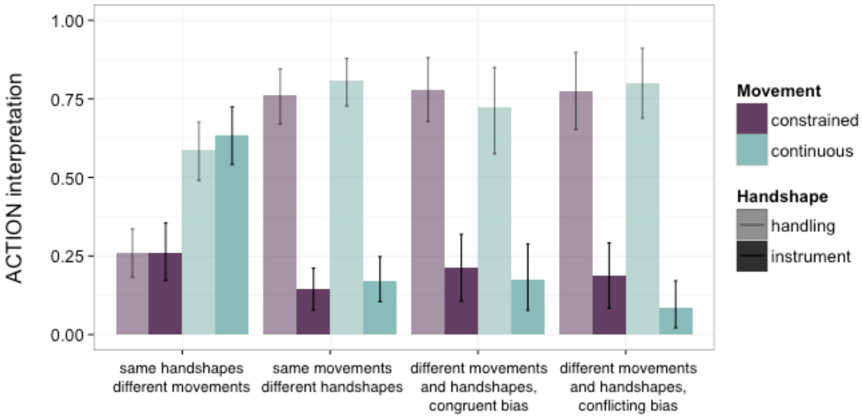


Figure 1. Proportion of gestures interpreted as describing an ACTION (as opposed to an OBJECT) for each of two gesture types, in each of the four experimental conditions

In this perception experiment, we replicate the finding that nonsigners exhibit a handshape bias when interpreting gestures as referring to objects/actions, and we identify an expected movement bias. When these biases are in conflict, nonsigners overwhelmingly map forms to meanings on the basis of their handshape bias.

References

Croft, W. (2000). Parts of speech as language universals and as language-particular categories. In P.M. Vogel and B. Comrie (Eds.), *Approaches to the Typology of Word Classes* (pp. 65-102). New York: Mouton de Gruyter.

Johnston, T. (2001). Nouns and verbs in Australian sign language: An open or shut case? *Journal of Deaf Studies and Deaf Education*, 6(4), 235-257.

Langacker, R. W. (1987) Nouns and verbs. *Language* 63(1), 53-94.

Padden, C., Hwang, S.-O., Lepic, R., & Seegers, S. (2015). Tools for language: Patterned iconicity in sign language nouns and verbs. *TiCS* 7, 81-94.

Supalla, T. & Newport, E. (1978). How many sits in a chair? The derivation of nouns and verbs in American Sign Language. In P. Siple (Ed.), *Understanding language through sign language research* (pp. 91-132). New York: Academic Press.

Tkachman, O. & Sandler, W. (2013) The noun-verb distinction in two young sign languages. In Haviland, John B. (ed.), *Where Do Nouns Come From?* Gesture 13:3, 147–180.

Verhoef, T., Padden, C. and Kirby, S. (2016). Iconicity, Naturalness And Systematicity In The Emergence Of Sign Language Structure. In S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér & T. Verhoef (eds.) *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11).* http://evolang.org/neworleans/papers/47.html

DID LANGUAGE EVOLVE FROM INDEXICAL SIGNALING?

TILL NIKOLAUS VON HEISELER*[1],

*Corresponding Author: formatlabor.net@gmail.com
[1]Department of Philosophy, Humboldt-Universität zu Berlin, Germany

This paper discusses indexical signaling as a possible precursor of declarative sentences in the evolution of language. Such a proposal is based on two assumptions. The first is known as the social intelligence hypothesis. The second is the function-first approach to explaining the evolution of traits: before a prototype of a new trait develops and the adaptation process begins, something already existing is used for a new purpose. Applied to the emergence of declarative sentences, this suggests that for language evolution to begin, something already existing was used for a declarative function (expressing a proposition). Thus, the evolutionary hypothesis presented here is that before human language began to develop, natural signs (such as indexical objects) were integrated into communication. I show that such a behavioral display can imply a conceptual structure similar to that informing the syntax of sentences: the displayer represents the thematic role of agent, while an indexical object (e.g. a hunting trophy) plays the role of patient.

1. Introduction

There are, roughly, two views on language which have dominated the debate about language evolution in the last few decades. (1) Traditionally, theorists of "generative grammar" conceptualized language as a productive system, in which a finite number of elements can generate a theoretically infinite number of sentences (Chomsky, 1965). (2) More recently, theories of embodied cognition claim that higher cognitive functions, including language, are rooted in lower cognitive functions, such as the sensory-motor system (Arbib, 2005).

Though both theories of language are valid and give valuable insights, they have their shortcomings regarding language evolution. The theory of generative grammar has resisted explanation in terms of evolution, mainly because it conceptualizes language as an internal symbolic machine that either works or does not (Hauser, Chomsky, & Fitch, 2002). The embodied cognition approach, on the other hand, succeeds in identifying some requirements of linguistic communication and gives many insights into the situatedness of linguistic communication; however, it has not yet explained the development of *syntax* and the emergence of the truth value (as a binary cognitive structure underlying the interpretation of declarative sentences) in terms of evolution.

In this paper I show how the idea of embodiment can address the problem of the emergence of syntax by including natural (indexical) signals in communication, and how, as a second step, the display of an object with indexical meaning could spur the development of syntax-like conceptual structures in which fundamental thematic roles (such as agent and patient) are marked by the direction of a mimetic gesture signifying the verb. For the sake of simplicity, I shall propose my hypothesis in form of statements and refrain from hedging.

2. Three problems in language evolution

There are at least three problems concerning language evolution. The first two are fundamental to any evolutionary history and originally addressed by Tinbergen (1963)—questions about the two "ultimate causes," function and evolutionary process.

(1) *The function*. What did the trait evolve for? Here the challenge is *to suggest the right kind of selective pressure*. Language, for instance, could not have evolved for transmitting valuable information, because this would give a reproductive advantage primarily to the receiver (and not to the sender). Also, language could not have evolved for better coordination, because this would give an advantage to the group, and no complex trait can evolve by group selection (Williams, 1966). In an ideal evolutionary scenario, a slightly better speaker would out-reproduce all others.

(2) *The evolutionary process*. How did the trait develop *gradually*? This relates to the challenge of *irreducible complexity*. Syntax needs a minimal complexity to fulfill its function, for instance to refer to an absent action. Also, typical declarations imply truth values, which suggests that declarative sentences and truth values emerged simultaneously.

(3) The third problem is specific to language. When we use language, we cooperate. For instance, we mostly speak the truth and assume that other people do so as well (Grice, 1975). If people didn't follow this *cooperative principle* linguistic communication would fail. Therefore, many researchers (Ferretti et al., 2017) have said that, before language could even begin to evolve, a kind of cooperative principle had to be in place. On the other hand, researchers who work on the problem of cooperative behavior suggest that human altruism is based on gossip (Barclay, 2010). If gossip is circulating in a group and reputation is based on what other people say about you, then it might be rational to develop cooperative behavior. This obviously creates a paradox.

3. The evolutionary turning point

We define the evolutionary turning point as the point at which a function is fulfilled for the first time, creating a reproductive advantage. The idea is that, before a trait can develop, something already existing is used for a new purpose (the function the trait will later develop to fulfill). This can be illustrated by the development of tetrapods. Walking on four feet developed successfully only once. All land animals—including reptiles, birds, and mammals—descended from one kind of bony fish. The crucial point in the evolutionary development of four-footed ambulation was not the movement from water to land but the point at which an individual could improve its reproductive success by using its fins as legs in the water to walk on the bottom of the ocean. After this turning point, the trait developed depending on various factors including selective pressures, genetic variation in the population, the frequency of occurrence of mutations beneficial for the trait and the size of the population. However, we should remember that this process began with some individuals out-producing all others by being better walkers. Here, previously existing structures (the fins) are used for a new function (walking).

4. What is the evolutionary function of language?

The evolutionary function of a trait is the function it evolves for. Textbooks often classify sentences by function: exclamations, imperatives, declarations and questions. We find the exclamative function in most mammals and birds, while the gestures of chimpanzees fulfill a imperative function (Hobaiter & Byrne, 2014). Questions imply declarations as answers. This might suggest that language developed, primarily, for the function declarations can fulfill. Most theorists agree that declarative sentences fulfill the function of transmitting propositions (Akmajian, 1984). If this is true, the original evolutionary function of language can be specified as the simplest form of such transmission.

5. What is the cognitively simplest declaration?

Cognitive simplicity is a relation between a cognitive system and a structure to be understood; this makes it sometimes distinct from technical or logical simplicity. The technically simplest syntactic structure that can transmit declarative content, including the representation of displaced action, consists of one verb and one argument, representing the subject. However, the mirror neuron system of primates encodes only *transitive* actions (actions with objects; Rizzolatti & Arbib, 1998). Therefore, the cognitively simplest syntactic structure corresponding to a pre-linguistic primate conceptual structure includes not only a verb and an actor, but also a *patient* (or *theme*).

Accepting this argument suggests: (1) that the simplest transitive verbs are *bivalent* with two semantic argument slots (agent, patient).

(2) Verbs that express *observable actions* are cognitively simpler than abstract verbs. "Jill hits Jack" is simpler than "Jill despises Jack." The simplest propositions describing observable actions also always use *singular terms* (terms referring to concrete objects).

(3) The simplest *aspect* (how actions, states or events extend over time) is a *single* event that happened at one particular point in time.

(4) The simplest and easiest kind of non-present action to represent is one still fresh in memory. In contrast, to draw attention to an overlooked present would require only an indexical call (such as an alarm or food call), which are part of animal communication. In other words, only reference to some non-present actions requires the essential features of human language (such as syntax and symbols), and therefore these constitute the simplest relevant scenarios.

(5) The simplest grammatical person to express seems to be the first-person singular, which is always present and often implied. In sign languages first-person singular sentences can be expressed without role-taking (while other grammatical persons require perspective switches between the verb—signed by the speaker—and the grammatical person; Janzen, 2017).

(6) Mimetic gestures are less demanding than conventional ones.

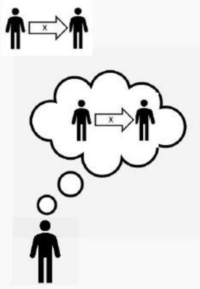
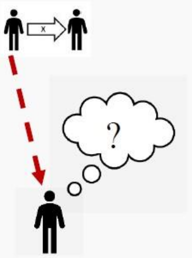
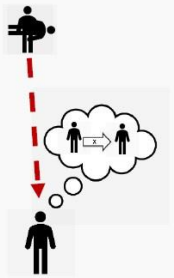
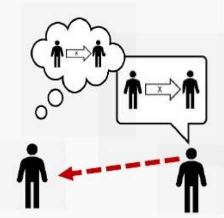
(7) In most sign languages, a sentence can be expressed by including the sender and the receiver and other present objects. "I give you the book" can be expressed with only one gesture directed from *me* (the speaker) to *you* (the receiver). The *giver* and the *receiver* are marked by the *direction* of the gesture.

Arguments 1–7 can be brought together in the following way: The simplest declarative sentence refers to a single concrete past transitive action consisting of three elements: a) the sender as the agent, b) a present object as the patient, c) a directed mimetic sign—representing the action—that marks the agent and the patient in their semantic roles.

6. Declarations before language

The question arises of what could be used to refer one's own past action, when there is no language: no symbol use, no understanding mimetic objects or gestures? At the turning point of language evolution, something which already existed must have been used for transmitting information about a concrete past action of the sender, with results beneficial to the reproduction of the sender's genes. Anything purposely used to refer to a non-present action must be considered some kind of sign. There are two classes of signs: natural and non-natural signs (Grice, 1957). The latter include mimetic or conventional symbols and are expressed intentionally. Natural signs, in contrast, are causally related to what they designate. All natural signs are indexical. Consequently, anything that can refer to the past of the sender in a non-linguistic world is an indexical sign. For an evolutionary process to begin, the transmission of this information needs to be beneficial to the sender. Following the function-first approach to explaining evolution (von Heiseler, 2019), these natural indexes referring to past actions of the sender would be integrated into communication and this communication would develop into language.

Table 1: *The three stages of transmitting information about one's past*. A = sender, B = patient, C = receiver. The interrupted arrow signifies the transmission of the declarative content. Left column: primate cognition; the two middle columns show two different scenarios of indexical signaling; in column 2 the patient is a living creature, in column 3 an inanimate object.

(1) Action comprehension	(2) Demonstrative behavior	(3) Trophy presentation	(4) Linguistic utterance
			
Inner representation of transitive actions. C understands what A does to B when (s)he sees it.	A shows C what (s)he did to B by restarting action x when C appears. The information (not necessarily the understanding) transmitted from A to C has what A did to B as content.	A displays B as a war trophy to C. C *might develop the understanding* of what A did to B. Displaying a war trophy is comparable to saying "I killed B."	Linguistic utterances which can represent non-present actions: A makes a statement to C about what (s)he did to B. C understands what A did to B.
Some primates (Pellegrino, et al., 1992).	Some monkeys, apes and hominins*** (Cheney & Seyfarth, 2008).	Some apes; carrying trophies without display (Carvalho, et al., 2010; Hirata, et al., 2001); hominins developing display behavior.	Some humans (such as sapiens and Neanderthals, probably Homo erectus; Everett, 2017).

One possible categorization of indexical signaling follows the classification of the patients: the indexical sign can be either another animal, for instance a conspecific (see table 1, column 2) or an indexical object, including but not limited to war or hunting trophies (see table 1, column 3). We define *a trophy as any object that* *transmits information about the past of the displayer which is beneficial to the displayer.*

The elementary semantic function—the distinction between representation and reality—is a necessary prerequisite for full interpretation of a declaration. The simplest, and almost surely the first, way that our ancestors came to understand the distinction between representation and reality must be in situations where another individual expresses a contradictory belief. For this a belief has to be deniable. This rarely is the case in situations in which the indexical object of the action is a living creature: if, for instance, A defeats B, C arrives, and A hits B again to inform C, B can react in two ways: either (s)he can retreat, in which case it is true that A has defeated B, or (s)he can fight back, in which case the fight is not yet settled. This is to say: the reaction of the patient of the action is a reliable truthmaker.

Furthermore, a threatening gesture directed at one individual with a communicative function to another individual is a behavior we find in apes and monkeys (Cheney & Seyfarth, 2008), without them developing language-like behavior. All these objections make it unlikely that language evolved from this kind of demonstrative behavior. Let us therefore look again at the trophy display behavior.

7. From trophy display to simple language

Trophy display has a syntax-like structure (figure 1a). When individual A presents the head of a lion, this can be interpreted as the declaration “I killed this lion.” The displayer signifies the agent, the trophy incorporates the patient, and the state of the trophy implies the concept of “killing.”

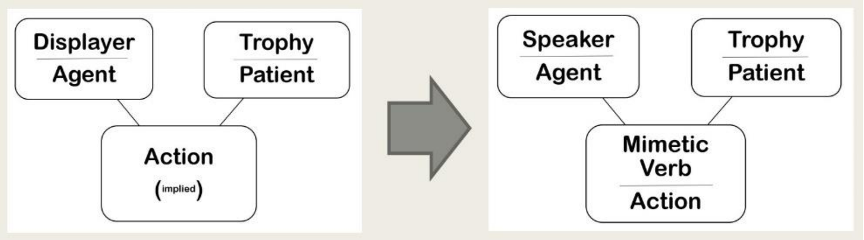


Figure 1: *The transition to language*. Left: the structure of trophy display (1a); right: the structure of the simplest sentence (1b; cf. section 5).

If, now, A adds a directed mimetic gesture—signifying the verb (“killing”)—this makes the trophy display behavior convey the same kind of information as a cognitively simplest sentence (as argued earlier). Here the syntax is expressed as the relation between present objects marked by the direction of the mimetic gesture. The gesture signifying the verb is in this case directed from the agent to the patient.

8. Solving the three problems

(1) We presented a scenario in which the speaker gains a reproductive advantage. Since speakers talk about themselves, they could influence the mental states of others in a beneficial way, which might influence social hierarchy or sexual selection.

(2) Irreducible complexity is not a problem. The step from the trophy display to a simple declarative sentence is rather small: You only need to add one directed mimetic gesture as a verb.

(3) The *cooperative principle* does not need to be in place. The trophy indicates the past action of the speaker; though it can be doubted, declarations including trophies bear costs.

It might be rewarding to further explore in what kind of contexts the mimetic gesture could have first emerged. One possibility is that the mimetic gesture could be made as a reaction to a sign on the part of the receiver that they do not believe the implications of the trophy display (e.g. that A made the kill). This would make the first declaration a negation of a negation of an implicit statement of the trophy display. I suggest that the understanding of binary truth-values underlying any declaration is of social origin, founded on disagreements based on conflicts of implicit interests. I therefore deny that the concept of truth could have emerged through self-questioning about what one ought to believe, without any social interactions. In the latter case, the propositional attitude would develop as self-reflection. To attack this Cartesian view is the main target of this essay.

References

Akmajian, A. (1984). Sentence types and the form-function fit. *Natural Language & Linguistic Theory.*

Arbib, M. A. (2005). From money-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences,* 28, pp. 105–167.

Ardila, A. (2016). The Evolutionary Concept of “Preadaptation” Applied to Cognitive Neurosciences. *Frontiers in Neuroscience.*

Cheney, D. L., & Seyfarth, R. M. (2008). *Baboon Metaphysics: The Evolution of a Social Mind.* Chicago: University of Chicago Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge MA: MIT Press.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.

Ferretti, F., Adornetti, I., Chiera, A., Nicchiarelli, S., Magni, R., Valeri, G., & Marini, A. (2017). Mental Time Travel and language evolution: a narrative account of the origins of human communication. *Language Sciences*.

Gould, S. J., & Vrba, E. S. (1982). Exaptation – a missing term in the science of form. *Paleobiology,* 8 (1).

Grice, H. P. (1975). Logic and Conversation. In I. P. Cole, & J. L. Morgan (Eds.), *Speech acts.* (pp. 41-58). New York u. a.: Academic Press.

Grice, P. H. (1957). Meaning. *The Philosophical Review*, 66, pp. 377–388.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002, November 22). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, 298 no. 5598, pp. 1569–1579.

von Heiseler, T. N., 2019. Syntax of Testimony: Indexical Objects, Syntax, and Language or How to tell a story without words. *Frontiers in Psychology*.

Hobaiter, C., & Byrne, R. W. (2014). The Meanings of Chimpanzee Gestures. *Current Biology*.

Janzen, T. (2017). Composite utterances in a signed language: Topic constructions and perspective-taking in ASL. *Cognitive Linguistics*.

Rizzolatti, G., & Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences*.

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift für Tierpsychologie*, 20, pp. 410–433.

Williams, G. (1966). *Adaptation and Natural Selection*. Princeton: Princeton University Press.

LANGUAGE-SPECIFIC CONSTRAINTS ON WORD FORM PREDICT SEGMENT INFORMATION ACROSS THE WORD

Andrew Wedel[1], Adam King[1], Adam Ussishkin[1]

*Corresponding Author: wedel@email.arizona.edu
[1]Linguistics Department, University of Arizona, Tucson, USA

King & Wedel (in press) found that each of 20 languages within a typologically-diverse set exhibited a significant inverse correlation between word probability and segmental information, that is, words which are less probable tend to contain segments that disambiguate from more lexical alternatives. They interpreted this in terms of Zipf's Law of Abbreviation (1949) which argued that the negative correlation between word probability and length arises from effort reduction on the one hand and maintainance of sufficient information in the signal on the other. In this case, King & Wedel argued that over time, greater phonetic reduction in higher probability words not only tends to promote loss of segments (Zipf 1949; Kanwal et al. 2017; Mahowald et al. 2018), but also leads to creation of more common segments and segment sequences, leading to relatively lower disambiguating information in the segments of high probablity words.

Lexical access in listeners proceeds incrementally as the speech stream is perceived, with the result that early segments in a word tend to provide more information than later segments (van Son & Pols 2003; Magnuson et al. 2007). On this basis, King & Wedel predicted that the correlation between word probability and segment information should evolve to be strongest at word beginnings, and decay late in the word. Using a measure of the difference in correlation between word probability and segment information early versus late in the word (see King & Wedel in press for method details), they found in fact that the majority of languages in the dataset did show the predicted pattern. However, a minority of languages, such as Hebrew and Arabic, did *not* show a significantly greater correlation early in the word. The question we address here is why some languages do not show this predicted pattern.

We test the hypothesis that the failure to show preferential optimization of segment information early in the word arises from language specific constraints on word formation which create a denser lexicon, that is, a lexicon in which words tend to be disambiguated from each other by fewer segments. As an example, the lexical meaning of words in Semitic languages like Hebrew and Arabic is largely carried by tri-consonantal roots. This restriction of word contrast to just three consonants means that for most words, all three consonants are required to disambiguate from alternatives (Ussishkin 2005). In contrast, languages like Georgian and Dutch have relatively large phoneme inventories, complex syllable structures, and allow variable word lengths, with the result that these lexicons tend to be sparser so later segments in the word tend to provide less information. Building on Wedel, Ussishkin & King (2019), we hypothesize that languages with sparser lexicons like Georgian and Dutch tolerate relatively greater reduction late in the word over time because those later segments are less likely to be informative to begin with, resulting in a pattern in which low probability words show their highest information segments at their beginnings. Conversely, when language specific constraints result in a more densely packed lexicon, segments across the word contribute more evenly to lexical disambiguation and so lower probability words show more evenly high segment information across the word.

In this study we use overall mean edit distance between words as a proxy measure for average lexicon density. As described above, King & Wedel showed that for most languages in the dataset, there was a significant interaction between word probability and a factor measuring bias of higher segment information toward the word-beginning, indicating that lower probability words tend to have a greater bias toward high early segment information. Here, we show that as hypothesized, mean edit distance itself significantly predicts the strength of this relationship, where low mean edit distance (e.g., as in Hebrew) is correlated with a lower bias toward early higher information (see Fig 1.) We have tested this relationship with a variety of other approaches and it remains robust. These findings contribute to the growing body of work in linguistic and cultural evolution on the influence of external constraints on the development of system-internal patterns.

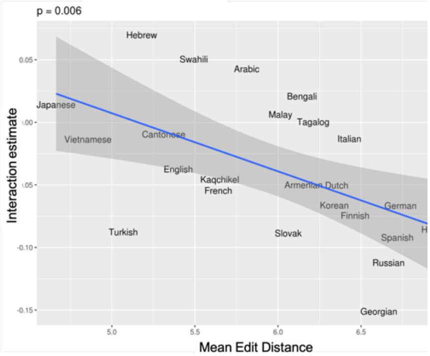


Figure 1. Correlation between mean edit distance and individual language model estimates for the interaction of word probability and early-bias for higher segment information. Languages with a higher mean edit distance (i.e, a sparser lexicon) are significantly more likely to show a bias toward early high segment information in lower probability words.

References

Kanwal, J., Smith, K., Culbertson, J., & Kirby, S. (2017). Zipf's law of abbreviation and the principle of least effort: Language users optimise a miniature lexicon for efficient communication. *Cognition*, 165, 45–52.

King, A. & Wedel, A. (in press). Greater early disambiguating information for less probable words: the lexicon is shaped by incremental processing. *Open Minds.*

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, 31(1), 133–156.

Mahowald, K., Dautriche, I., Gibson, E., & Piantadosi, S. T. (2018). Word forms are structured for efficient use. *Cognitive Science*, 42(8), 3116–3134.

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templaticmorphology. *Natural Language & Linguistic Theory*, 23(1), 169–218.

van Son, R., & Pols, L. C. (2003). How efficient is speech. In *Proceedings of the Institute of Phonetic Sciences*, 25, 171–184.

Wedel, A., Ussishkin, A., & King, A. (2019). Incremental word processing influences the evolution of phonotactic patterns. *Folia Linguistica*, 40(1), 231-248.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Welsey: Reading, Mass.

MODELLING THE ROLE OF OTHER-INITIATED REPAIR IN FACILITATING THE EMERGENCE OF COMPOSITIONALITY

MARIEKE WOENSDREGT*[1] and MARK DINGEMANSE[1]

*Corresponding Author: m.woensdregt@let.ru.nl
[1]Centre for language studies, Radboud University, Nijmegen, The Netherlands

A key feature setting apart human language from other animal communication systems is compositionality (Hockett, 1960). Recent work has focused on explaining the emergence of compositional language by reference to the combined pressures of learnability and expressivity (Kirby, Tamariz, Cornish, & Smith, 2015). Learnability alone would favour a degenerate system (few signals with the most generic meanings possible), while expressivity would be served equally well by a holistic or a compositional system. Only their interaction swings the balance in favour of compositionality, which provides the simplest (most compressible) way of expressing all meanings.

Here we connect these results to two other ubiquitous features of human communication: noise and interactive repair, or the metalinguistic use of signals to address trouble in producing, perceiving and understanding (Schegloff, Jefferson, & Sacks, 1977). Interactive repair has been shown to play a crucial role in streamlining social interaction in everyday and experimental settings (Micklos, Silva, & Fay, 2018; Fay et al., 2018). Cross-linguistic work has uncovered substantial pragmatic universals in this domain, including a basic division into OPEN vs. RESTRICTED repair initiation strategies and a drive to minimize joint effort in the collaborative resolution of trouble (Dingemanse et al., 2015). We aim to bring insights from this work to computational models of cultural evolution.

We adapt the model of Kirby et al. (2015) —in which languages are transmitted over generations through iterated Bayesian learning— and extend it to incorporate noise and repair under the following design decisions, all motivated by empirical observations of interaction. Speakers produce signals that are occasionally partially obscured by noise (Bergen & Goodman, 2015). Listeners may respond directly or initiate repair using either: an OPEN request (indicating no grasp of a signal's meaning and inviting full repetition) or a RESTRICTED request (indicating partial grasp and inviting partial repetition) (Dingemanse et al. 2015). Listeners initiate repair probabilistically depending on the level of ambiguity (uncertainty about the intended referent) and the cost associated with repair (Clark & Schaefer, 1987). This cost is higher for open than for restricted because the latter recycles part of the trouble source turn (Schegloff, 1979). We hypothesise that repair favours compositional systems under these assumptions:

1. A pressure for MINIMAL EFFORT (less costly responses are preferred)
2. A pressure to reach MUTUAL UNDERSTANDING (agents attempt to formulate interpretable signals and reach unambiguous interpretations)

We hypothesise that having only a pressure for minimal effort does not favour any particular language type, while a pressure for mutual understanding on its own will favour holistic and compositional languages equally. We hypothesise that their joint effect given the availability of repair favours compositional over holistic languages because compositional languages enable taking full advantage of the partial interpretability of signals. Learners of compositional languages should receive more reliable data in this condition, because the agents that produce the data are more likely to use repair to solve ambiguity caused by noise.

We find that the hypothesised effect of the combined pressures for mutual understanding and minimal effort appears when the probability of noise exceeds 0.5. When this is the case, the proportion of compositional languages relative to holistic languages increases when both pressures are combined, compared to when only a pressure for mutual understanding is present (Figure 1). (See supplementary materials for code and parameter space exploration.)

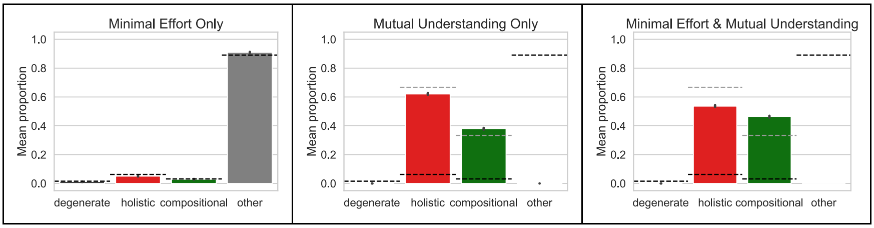


Figure 1. Mean proportions and 95% CIs of language types after convergence (generations 750-1,000 of 100 simulation runs). Dashed lines show baseline proportions of language types in hypothesis space (black for all types, grey for relative proportions of only the fully expressive language types). P(noise) = 0.6, cost ratio OPEN:RESTRICTED = 2:1, transmission bottleneck = 20 data points per learner.

In sum, we show that in addition to serving the combined pressures of learnability and expressivity, compositional languages can also be useful under the interactional dynamics of noise and repair, where they enable agents to reach mutual understanding with minimal effort.

Acknowledgements

MW and MD are grateful for funding from the Dutch Organisation of Scientific Research (NWO) grant 016.vidi.185.205 on 'Elementary particles of conversation'.

References

Bergen, L., & Goodman, N. D. (2015). The Strategic Use of Noise in Pragmatic Reasoning. *Topics in Cognitive Science*, *7*(2), 336–350. doi: 10.1111/tops.12144

Clark, H. H., & Schaefer, E. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, *2*(1), 19–41. doi: 10.1080/01690968708406350

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., … Enfield, N. J. (2015). Universal Principles in the Repair of Communication Problems. *PLOS ONE*, *10*(9), e0136100. doi: 10.1371/journal.pone.0136100

Fay, N., Walker, B., Swoboda, N., Umata, I., Fukaya, T., Katagiri, Y., & Garrod, S. (2018). Universal Principles of Human Communication: Preliminary Evidence From a Cross-cultural Communication Game. *Cognitive Science*, *42*(7), 2397–2413. doi: 10.1111/cogs.12664

Hockett, C. F. (1960). The Origin of Speech. *Scientific American*, *203*(3), 89–96.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102. doi: 10.1016/j.cognition.2015.03.016

Micklos, A., Silva, V. M., & Fay, N. (2018). The prevalence of repair in studies of language evolution. In C. Cuskley, M. Flaherty, H. Little, L. McCrohon, A. Ravignani, & T. Verhoef (Eds.), *The Evolution of Language: Proceedings of the 12th International Conference (EVOLANGXII)*. Online at urlhttp://evolang.org/torun/proceedings/papertemplate.html?p=136.

Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation. In T. Givón (Ed.), *Syntax and Semantics* (Vol. 12, pp. 261–286).

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The Preference for Self-Correction in the Organization of Repair in Conversation. *Language*, *53*(2), 361–382.

THAT'S (NOT) ABOUT THE SIZE OF IT: SOUND-SYMBOLIC EFFECT OF LABELS ON SIZE PERCEPTION

YA-NING WU[1], and ARTURS SEMENUKS[2]

Corresponding authors: asemenuk@ucsd.edu, yanniwu5523@gmail.com
[1]Institute of Cognitive Neuroscience, National Central University, Taoyuan, Taiwan,
[2]Department of Cognitive Science, University of California San Diego, La Jolla, USA

Sound symbolism is a set of cross-culturally common iconic associations between perceptual features, such as size, roundness or weight, and phonetic features of speech sounds (Sidhu & Pexman, 2018). A growing number of papers provides support to the notion that sound symbolism plays a role in the dynamics of language change on time scales from developmental (Imai et al., 2008) to cultural (Carr et al., 2018) to evolutionary (Cuskley & Kirby, 2013). Research suggests multiple functions it could serve and have served in language evolution, including a rather fundamental one – bootstrapping the emergence of language in the human lineage. Specifically, Cuskley and Kirby (2013) hypothesize that sound symbolism could have provided a foundational resource for the emergence of protolanguage via iconic cross-modal associations that were used for communication.

Thus, understanding the mechanics of sound symbolism is important for theories of language evolution. However, many questions remain understudied, including the question of the mechanism: what underlies the iconic mappings in sound symbolism? A frequently assumed answer is that sound symbolism works broadly like other cross-modal correspondences, e.g. between (non-vowel) pitch and size (Ramachandran & Hubbard, 2001). As cross-modal correspondences are not unique to humans, a common mechanism for the two types of phenomena would make the sound-symbolic bootstrapping of (proto)language hypothesis evolutionarily plausible (Cuskley & Kirby, 2013). The similarities between the phenomena make this assumption justifiable, but it cannot be blindly taken for granted, as important differences exist as well. For example, sound-symbolic mappings involve more complex features (e.g. vowel height) compared to other cross-modal correspondences (Parise, 2016). If sound-symbolic relationships are not bidirectional, the idea that they work like other types of cross-modal correspondences needs to be reexamined or further fleshed out. That, in turn, would affect the sound-symbolic bootstrapping hypothesis and other theories of the role of sound symbolism in language evolution that rely on this assumption. In order to shed more light on this question, we focus on bidirectionality of sound-symbolic mappings, as it is typically assumed that cross-modal correspondences are bidirectional (Deroy & Spence, 2013), but whether that is the case for sound symbolism has not been tested.

We investigated this question experimentally by testing whether people would more often misremember shapes as being smaller when they were given a nonword with a front vowel (/i/), as compared to a back vowel (/u/) in a forced choice task. We focused on this sound-symbolic mapping due to extensive research on the effect of vowel frontness and height on size perception (Sapir, 1929; Newman, 1933; Thompson & Estes, 2011; Knoeferle et al, 2017). Each shape was once paired with an /i/-nonword and once with an /u/-nonword. In each experimental trial, a shape accompanied by a nonword with CVC structure appeared on the screen for 5 seconds. Participants were instructed to remember the name of the shape. Following a 5 second presentation of a fixation cross, the participants were asked to type in the name of the shape without any feedback, after which they were presented with two shapes – one smaller than the original and one larger – and were asked to choose which shape they saw earlier.

Analyses of data from 63 participants (45 female, mean age = 20.03, SD = 1.67) using mixed effects logistic regression show a significant effect of vowel on shape size perception ($\chi 2 = 3.8635$, $p < 0.05$): participants chose the smaller picture 44.95% of the time when the nonword contained /u/, and 50.13% of the time when it contained /i/ (Fig. 1). The result thus supports the bidirectionality of sound-symbolic associations. That, in turn, supports the evolutionary plausibility of the role of sound symbolism in early stages of language evolution by contributing evidence for the continuity between sound-symbolic and other cross-modal associations.

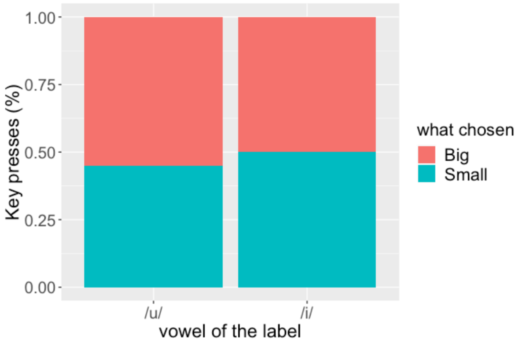


Fig 1. X-axis indicates the nonword vowel; the y-axis shows the percentage of choices for all participants. The color indicates which size choice was chosen by participants.

References

Carr, J. W., Smith, K., Cornish, H., & Kirby, S. (2017). The cultural evolution of structured languages in an open-ended, continuous world. *Cognitive science*, 41(4), 892-923.

Cuskley, C. , & Kirby, S. (2013) Synaesthesia, cross-modality and language evolution. In J. Simner & E. M. Hubbard (Eds.), *Oxford Handbook of Synaesthesia* (pp.869–907). Oxford: Oxford University Press.

Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic bulletin & review*, 20(4), 643-664.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, 109(1), 54-65.

Knoeferle, K., Li, J., Maggioni, E., & Spence, C. (2017). What drives sound symbolism? Different acoustic cues underlie sound-size and sound-shape mappings. *Scientific reports*, 7(1), 5562.

Newman, S. S. (1933). Further experiments in phonetic symbolism. *The American Journal of Psychology*, 45, 53–75. doi: 10.2307/1414186

Parise, C. V. (2016). Crossmodal correspondences: Standing issues and experimental guidelines. Multisensory research, 29(1-3), 7-28.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. Journal of Consciousness Studies, 8(12), 3-34.

Sapir, Edward. (1929). A study in phonetic symbolism. Journal of experimental psychology 12(3). 225-239.

Sidhu, D. M. & Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psychological Bulletin & Review* 25, 1619–1643. doi:10.3758/s13423-017-1361-1

Thompson, P. D., & Estes, Z. (2011). Sound symbolic naming of novel objects is a graded function. Quarterly Journal of Experimental Psychology, 64(12), 2392-2404.

YOUR THEORY OF LANGUAGE EVOLUTION ALSO DEPENDS ON YOUR VIEW OF EVOLUTION

THIAGO MACEK GONÇALVES ZAHN*[1]

*Corresponding Author: thimacek@gmail.com
[1]Departamento de Linguística, Universidade de São Paulo, São Paulo, Brazil

One exciting but often puzzling aspect of the field of Language Evolution is the variety of theories and approaches it includes. These not only embrace different, sometimes contradictory assumptions, but also often seek explanations to different questions. An interesting topic is thus how different theories/ approaches are related, i.e. how different assumptions may lead to disparate questions and views, and how contradictory different assumptions made in the field in fact are.

In an important contribution, Jackendoff (2010) argued that "what there is for a theory of language to explain" depends on one's theory of "what language is", discussing how different views on the "innate language capacity" and on how domain-specific it is lead to different theories on how that language capacity might have evolved. As Jackendoff mentions, some defend little or nothing special (i.e. domain-specific) is needed for the evolution of language, making it essentially a "cultural phenomenon". In the same article, however, he states that "if that ("disparate languages" passed down through cultural transmission) is all there is to language, a theory of the evolution of language has nothing at all to explain". Therefore, although Jackendoff speaks of cultural evolution, he sees a clear contrast between this and language evolution, apparently equated with biological evolution and deemed more important.

Others distinguish even more clearly between *evolution*, as a specifically biological phenomenon, and *language change*: Berwick & Chomsky (2016, p. 92) explicitly say "Languages change, but they do not evolve. (…) nonbiological evolution (…) is not evolution at all"; and Andersen (2006) has a whole chapter arguing "that there is no chance of explaining language *change* by the mechanisms of *evolutionary* theory". For some (e.g. Berwick & Chomsky), this distinction is taken as given - perhaps reflecting the Chomskyan view that linguistic variation and change are strongly constrained by a shared Universal Grammar. Others, like Andersen (2006) and Itkonen (1999), have discussed at some length what they see as *disanalogies* that would disavow the use of *evolutionary thinking* to explain *cultural* linguistic phenomena.

Although researchers who advocate for a cultural evolutionary approach to language change and *evolution* (i.e. *origin of modern human languages*) have provided some answers to such criticisms (e.g. Dediu et al, 2013, p. 305-307; Steels, 2017), little conversation seems to occur between both sides of the divide. In part, this may reflect different conceptions of *language*, as pointed by Jackendoff (2010), and even mirror the Formalist-Functionalist divide common in linguistics. However, it seems clear that another aspect underlying the divide are different understandings of *evolution* - including *what counts as evolution*, *what evolution can/ should explain* and related questions. This suggests that some of the ongoing debates in the field are *inherently theoretical*, and thus cannot be solved solely empirically, since new evidence may be interpreted differently and/ or given different weights depending on one's assumptions.

Discussion on "what is evolution" and on whether it is possible (or productive) to expand evolutionary thinking beyond biology has a prolific history in the philosophy of biology (e.g. Lewontin, 1970; Hull, 1988; Godfrey-Smith, 2007, 2009, 2012; also Price, 1995[1971]; Frank, 2012; Luque, 2017). Assuming an ontology based on a 'general selection theory' (e.g. Croft, 2000; Clark, 2010; Gong, 2012; Steels & Szathmáry, 2018), 'units of language' may be interpreted as 'units of selection' and/or linguistic *replicators* and *interactors* (*sensu* Hull, 1988) may be defined, allowing *linguistic* changes to be studied alongside or independently of genetic changes in speakers. In fact, models based on similar assumptions have shown how *cultural language evolution* could have shaped modern human language(s) (see e.g. Steels, 2010; Kirby, 2013). On the other hand, assuming that 'evolution' is limited to *genetic* changes raises questions about what is inherently different in *cultural* changes, and whether phenomena at that level might not have long-term effects in the evolution of human language(s).

Taking as a "general model" Kirby (2017, p.125)'s idea that language involves the interaction of three dynamical systems (*individual learning, cultural evolution and biological evolution*), important differences between theories may be recognized by considering which of these systems in fact involve *evolution*, which are deemed more important, and how they influence each other in each approach.

It is perhaps not surprising that one's theory of *Language Evolution* depends on one's theories of *language* and *evolution*; thus, clearing assumptions in both regards may go a great way in building more constructive exchanges in the field.

Acknowledgments

The author would like to thank João Figueiredo Nobre Cortese, Thomas Finbow, Paulo Chagas de Souza, Joana Bortolini Franco, Wellington Santos da Silva, Mariana Payno and Vitória Lima for constructive discussions about the nature of language and language evolution, and CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for financial support which was essential to allow the development of this research.

References

Andersen, H. (2006) Synchrony, Diachrony, and Evolution. In: O.N. Thomsen (Ed.) *Competing models of linguistic change – Evolution and Beyond* (pp. 59-90). Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Berwick, R., & Chomsky, N. (2016). Why only Us: Language and evolution. Cambridge, MA: MIT Press.

Clark, B. (2010). Evolutionary Frameworks for Language Change: The Price Equation Approach. *Language and Linguistics Compass 4(6)*: 363-376.

Croft, W. (2000). *Explaining Language Change: An Evolutionary Approach*. London: Longman.

Dediu, D., Cysouw, M., Levinson, S.C., Baronchelli, A., Christiansen, M.H., Croft, W., Evans, N., Garrod, S., Gray, R.D., Kandler, A., Lieven, E. (2013) Cultural Evolution of Language. In: P.J. Richerson and M.H. Christiansen (Eds.) *Cultural Evolution: Society, Technology, Language, and Religion*. (pp. 303-332). Cambridge, MA: MIT Press.

Frank, S.A. (2012). Natural selection. IV. The Price equation. *Journal of Evolutionary Biology* 25: 1002-1019.

Godfrey-Smith, P. (2007). Conditions for Evolution by Natural Selection. *Journal of Philosophy* 104: 478-516.

Godfrey-Smith, P. (2009). *Darwinian populations and natural selection*. New York, NY: Oxford UP.

Godfrey-Smith, P. (2012). Darwinism and cultural change. *Philosophical Transactions of the Royal Society B 367*: 2160-2170.

Gong, T., Shuai, L., Tamariz, M., Jäger, G. (2012). Studying Language Change Using Price Equation and Pólya-urn Dynamics. *PLoS ONE 7(3)*: e33171.

Hull, D.L. (1988). A mechanism and its metaphysics: An evolutionary account of the social and conceptual development of science. *Biology and Philosophy 3(2)*: 123-155.

Itkonen, E. (1999). Functionalism yes, biologism no. *Zeitschrift für Sprachwissenschaft*, *18(2)*: 219-221.

Jackendoff, R. (2010). Your theory of language evolution depends on your theory of language. In: R.K. Larson, V. Déprez, H. Yamakido (Eds.) *The Evolution* *of Human Language – Biolinguistic Perspectives* (pp. 63-72). Cambridge: Cambridge University Press.

Kirby, S. (2013). Transitions: The Evolution of Linguistic Replicators. In: Binde, P.-M.; Smith, K. (Eds). *The Language Phenomenon: Human Communication from Milliseconds to Millennia*. (pp. 121-138). Heidelberg: Springer.

Kirby, S. (2017). Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin and Review 24*: 118-137.

Lewontin, R.C. (1970). The units of selection. *Annual Review of Ecology and Systematics 1*: 1-18.

Luque, V.J. (2017). One equation to rule them all: a philosophical analysis of the Price equation. *Biology & Philosophy 32(1)*: 97-125.

Price, G.R. (1995[1971]). The Nature of Selection. *Journal of Theoretical Biology* 175: 389-396.

Steels, L. (2010). Can Evolutionary Linguistics Become a Science? *Journal for Evolutionary Linguistics* 1(1): 1-34.

Steels, L. (2017). Do languages evolve or merely change? *Journal of Neurolinguistics 43(B)*: 199-203.

Steels, L., Szathmáry, E. (2018). The evolutionary dynamics of language. *Biosystems 164*: 128-137.

EMERGENCE OF PRAGMATIC REASONING FROM LEAST-EFFORT OPTIMIZATION

NOGA ZASLAVSKY*[1], JENNIFER HU[1], and ROGER LEVY[1]

*Corresponding Author: nogazs@mit.edu
[1]Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, USA

It is widely acknowledged that pragmatics is an important driver of language evolution (e.g., Sperber & Origgi, 2010; Scott-Phillips, 2015; Woensdregt & Smith, 2018) and language change (Jucker, 1995; Traugott, 2012; inter alia). A sound theoretical understanding of the dynamics of pragmatic reasoning and computational principles that may give rise to this dynamics is thus crucial to the field of language evolution. Here, we study this dynamics within the Rational Speech Act framework (RSA: Frank & Goodman, 2012; Goodman & Frank, 2016).

RSA formulates pragmatic reasoning as probabilistic speakers and listeners recursively reasoning about each other. The speaker is defined by a production distribution $S(u|m)$ over possible utterances u given meaning m, and the listener is defined by an inference distribution $L(m|u)$. RSA recursively relates the speaker and listener by assuming a Bayesian listener—$L(m|u) \propto S(u|m)P(m)$, with $P(m)$ a prior distribution on speaker meanings that is assumed to be in common ground—and a speaker that is bounded-rational with respect to a utility function $V(u, m)$ (typically, $V(m, u) = \log L(m|u) - C(u)$ where $C(u)$ specifies the cost of u). That is, $S(u|m) \propto \exp(\alpha V(u, m))$, where α controls the degree to which the speaker maximizes utility. The framework enjoys broad popularity and empirical support (for review: Goodman & Frank, 2016), and while shallow recursion is often assumed, several studies have also explored and motivated deeper recursions (e.g., Camerer, Ho, & Chong, 2004; Franke & Degen, 2016; Bergen, Levy, & Goodman, 2016; Levy, 2018). These explorations have relied on numeric simulation (e.g., Yuan, Monroe, Bai, & Kushman, 2018; Peloquin, Goodman, & Frank, 2019), leaving much unknown regarding the dynamics of RSA recursion.

Here we present new analytic results, illustrated by implemented model instances, that answer key open questions about RSA dynamics. Because the RSA speaker is guided by (soft) optimization of utterance utility, the intuition is widely held that RSA recursion is guaranteed to (locally) optimize expected utility (e.g., Yuan et al., 2018). Our analysis disconfirms this intuition. We show that the RSA recursion is an instance of the alternating maximization algorithm (Csiszár & Shields, 2004), providing an optimization guarantee. However, the guarantee is

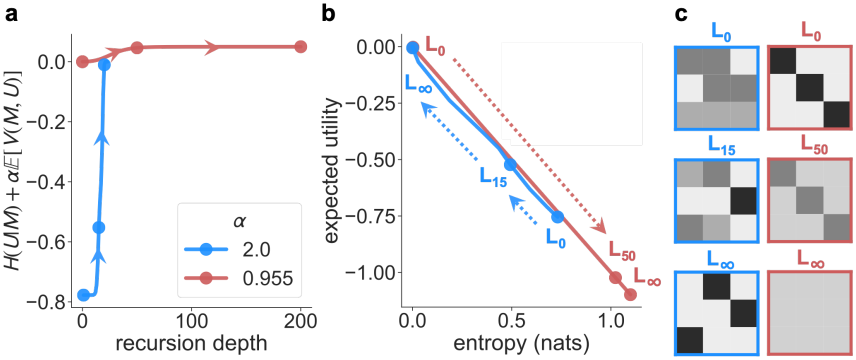


Figure 1. Model simulations with three uniformly distributed meanings and three possible utterances. **(a)** Trade-off between utility and effort improves with the depth of recursion. **(b)** Expected utility as a function of the speaker's entropy, $H(U|M)$. Expected utility may increase (blue; $\alpha = 2$) or decrease (red; $\alpha \approx 0.955$) as recursion depth increases. **(c)** Listener distributions at initial, intermediate, and converged conditions. Darker grays correspond to higher probabilities. (For L_0 in red, the off-diagonal elements are initialized to small but non-zero ϵ values.)

not to improve expected utility but rather a *tradeoff* between communicative effort and expected utility, namely $H(U|M) + \alpha\,\mathbb{E}[V(M,U)]$, where $H(U|M)$ is the conditional entropy of utterances given speaker meanings. This tradeoff can be thought of as an instance of Zipf's least-effort principle (Zipf, 1949), where here low communicative effort corresponds to high entropy of the speaker's production distribution. Our analysis also reveals that in general α does not simply trade off against recursion depth, as widely understood (e.g., Frank, Emilsson, Peloquin, Goodman, & Potts, 2018): the value of α determines the tradeoff between effort and communicative utility optimized by RSA recursion.

The model simulations of Figure 1 exemplify these results. RSA iteration always improves the utility–effort tradeoff (Figure 1a), but expected utility may increase (Figure 1b, blue trajectory), or decrease (red trajectory), depending on α and the initial listener (Figure 1c, L_0). We speculate that the possibility of RSA iteration decreasing expected utility has not previously been identified in numeric simulations because RSA initializations are typically (apart from structural zeroes arising when some messages do not satisfy the truth conditions of some utterances) already high in speaker conditional entropy $H(U|M)$.

This work shows that least–effort optimization, and not simply heuristic utility maximization, may give rise to human pragmatic reasoning. Therefore, the optimization principle we identified may shape the evolution of pragmatic skills and more generally, the evolution of language. In addition, we have directly linked the dynamics of RSA recursive reasoning to the dynamics of the known alternating minimization algorithm. This provides new theoretical grounds for further studying the dynamics of pragmatic reasoning and its role in language evolution.

References

Bergen, L., Levy, R., & Goodman, N. (2016). Pragmatic reasoning through semantic inference. *Semantics and Pragmatics*, *9*(20).

Camerer, C. F., Ho, T., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, *119*(3), 861-898.

Csiszár, I., & Shields, P. (2004). Information theory and statistics: A tutorial. *Foundations and Trends in Communications and Information Theory*, *1*(4), 417-528.

Frank, M. C., Emilsson, A. G., Peloquin, B., Goodman, N. D., & Potts, C. (2018). *Rational speech act models of pragmatic reasoning in reference games.* (PsyArXiv preprint of July 2, 2018, https://doi.org/10.31234/osf.io/f9y6b)

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998–998.

Franke, M., & Degen, J. (2016). Reasoning in reference games: Individual- vs. population-level probabilistic modeling. *PLOS ONE*, *11*(5), 1-25.

Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818-829.

Jucker, A. (Ed.). (1995). *Historical pragmatics: Pragmatic developments in the history of English* (Vol. 35). John Benjamins Publishing.

Levy, R. P. (2018). Communicative efficiency, uniform information density, and the rational speech act theory. In T. Rogers, M. Rau, X. Zhu, & C. Kalish (Eds.), *Proceedings of the 40th annual meeting of the cognitive science society* (pp. 684–689). Austin, TX: Cognitive Science Society.

Peloquin, B. N., Goodman, N. D., & Frank, M. C. (2019). The interactions of rational, pragmatic agents lead to efficient language structure and use. In *Proceedings of the 41st annual meeting of the Cognitive Science Society* (pp. 912–917). Austin, TX: Cognitive Science Society.

Scott-Phillips, T. C. (2015). Nonhuman primate communication, pragmatics, and the origins of language. *Current Anthropology*, *56*(1), 56-80.

Sperber, D., & Origgi, G. (2010). A pragmatic perspective on the evolution of language. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language: Biolinguistic perspectives* (p. 124–132). Cambridge University Press.

Traugott, E. C. (2012). Pragmatics and language change. In K. Allan & K. M. Jaszczolt (Eds.), *The Cambridge handbook of Pragmatics* (pp. 549–566). Cambridge University Press.

Woensdregt, M., & Smith, K. (2018). *Pragmatics and language evolution.* PsyArXiv.

Yuan, A., Monroe, W., Bai, Y., & Kushman, N. (2018). Understanding the Rational Speech Act model. In T. Rogers, M. Rau, X. Zhu, & C. Kalish (Eds.), *Proceedings of the 40th annual conference of the cognitive science society* (pp. 2759–2764). Austin, TX: Cognitive Science Society.

Zipf, G. K. (1949). *Human behavior and the principle of least effort.* Addison-Wesley (Reading MA).

Author Index

Özyürek, Asli, 356

Abels, Klaus, 263
Abner, Natasha, 296
Adger, David, 263
Alday, Phillip, 252
Andirkó, Alejandro, 294
Angwela, Felix O., 420
Anikin, Andrey, 186
Anton, Jean-Luc, 12
Arnon, Inbal, 437
Asano, Rie, 8

Ball, Gregory, 105
Barrett, H. Clark, 423
Becker, Yannick, 12
Berdicevskis, Aleksandrs, 14
Bergman, Thore J., 420
Beuls, Katrien, 311
Bickel, Balthasar, 181
Blaszczyk, Maryjka B., 420
Boë, Louis-Jean, 22
Boeckx, Cedric, 56, 266, 294, 324, 417
Boitelle, Pascaline, 12
Bolt, Laura M., 420
Briseño-Jaramillo, Margarita, 420
Brown, Michelle, 420
Carcassi, Fausto, 24
Carlson, Daryll, 59, 63
Cartmill, Erica A., 27, 423
Celma-Miralles, Alexandre, 371
Chen, Sihan, 30
Chen-Kraus, Chloe, 420
Christiansen, Morten H., 41, 125
Christopher, Jennie, 260
Claidière, Nicolas, 237, 377
Clark, Robin, 362
Clay, Zanna, 420
Collins, Michael, 38
Contreras Kallens, Pablo, 41
Cormier, Kearsy, 159, 390
Coupé, Christophe, 44
Courville, Aaron, 314
Coye, Camille, 420
Craig, Adrian, 155
Crasborn, Onno, 302
Cristea, Mioara, 334
Cross, Megan, 47
Culbertson, Jennifer, 197, 225, 263, 321
Cunningham, Sheila, 223
Curcic-Blake, Branislava, 162
Currie, Thomas, 177
Cuskley, Christine, 50

Davis, Charles, 53
Davis, Emily, 240

de Boer, Bart, 56, 63, 302
De Gregorio, Chiara, 447
de Heer Kloots, Marianne, 59
De Luca, Margherita, 53
de Reus, Koen, 59, 63
de Vos, Connie, 302
Deacon, Terrence, 93
Dediu, Dan, 44, 67, 426
Demolin, Didier, 71
Devanna, Paolo, 280
Dingemanse, Mark, 356, 474
Dooling, Robert, 105
Dumasy, Marie, 12

Eisen, Elad, 128
Emery Thompson, Melissa, 420
Emmorey, Karen, 74
Engesser, Sabrina, 81
Ergin, Rabia, 84
Essers, Charlotte, 87
Estrada, Alejandro, 420
Everett, Caleb, 30, 90

Fagot, Joël, 237, 291, 377, 451
Farias-Virgens, Madza, 93
Fay, Nicolas, 457
Feeney, Andrew, 95
Fenk-Oczlon, Gertraud, 98
Ferdinand, Vanessa, 102, 237
Ferrer-i-Cancho, Ramon, 125
Fichtel, Claudia, 420
Fikkert, Paula, 302
Firzlaff, Uwe, 280
Fishbein, Adam, 105
Flaherty, Molly, 108, 296
Fröhlich, Marlen, 122
Frank, Stella, 111, 131
Fruth, Barbara, 420
Fusaroli, Riccardo, 318, 457

Gómez, David Maximiliano, 331
Gómez-Maureira, Marcello A., 454
Gamba, Marco, 420, 447
Gao, Kexin, 197
Garcia, Maxime, 59, 63
Gauvrit, Nicolas, 223
Genter, Timothy, 374
Giacoma, Cristina, 420, 447
Giacomino, Laura, 12
Goldin-Meadow, Susan, 296
Gomez-Rodriguez, Carlos, 125
Graham, Kirsty E., 420
Green, Samantha, 420
Greenhill, Simon, 411
Groom, Nicholas, 174
Gross, Stephanie, 63
Grossman, Eitan, 128
Grueter, Cyril, 420
Gulamani, Sannah, 159
Guo, Shangmin, 131
Gupta, Shreejata, 420
Gustison, Morgan L., 420

Hörpel, Stephen, 280
Haas, Caroline, 63
Hagberg, Lindsey, 420
Hagoort, Peter, 165
Hartmann, Stefan, 141
Hausberger, Martine, 155
Havrylov, Serhii, 131
Hedwig, Daniela, 420
Heimann, Katrin, 457
Heintz, Christophe, 144
Henry, Laurence, 155
Hodge, Gabrielle, 159
Hoeksema, Nienke, 162, 165
Hosaka, Michio, 168
Hu, Jennifer, 484
Huerta-Sanchez, Emilia, 93
Hugget, Elliot, 171

Izawa, Ei-Ichi, 380

Jack, Grieve, 174
Jack, Katharine M., 420
Jadoul, Yannick, 63
Ji, Ting, 177
Jing, Yingqi, 181
Johannsen, Niels Nørkjær, 457
Johansson, Niklas, 186
Johansson, Sverker, 189
Johnson, Tamar, 197
Jon-And, Anna, 200
Jones, Keith, 209
Jowett, Adam, 405

Kanampiu, Patrick, 263
Kappeler, Peter M., 420
Kapron-King, Anna, 220
Kavanagh, Eithne, 420
Kemp, Charles, 102
Kempe, Vera, 223
Kiliaan, Amanda, 165
Kindellan, Rachel, 225
King, Adam, 234, 471
King-Bailey, Gillian, 420
Kirby, Simon, 24, 50, 108, 220, 237, 296, 321, 440
Kita, Sotaro, 81
Kluender, Robert, 240
Kotz, Sonja, 59
Kuběnová, Barbora, 420

Lacoste, Romain, 12
Lacroix, Travis, 314
Lameira, Adriano, 420
Lange, Léa, 155
Lattenkamp, Ella, 249
Lazaridou, Angeliki, 314
Lei, Li, 252
Lemasson, Alban, 420
Lepic, Ryan, 460
Leroux, Maël, 155, 255
Lev-Ari, Shiri, 257, 359
Levy, Roger, 484
Linnenschmidt, Meike, 249
Lombard, Marlize, 457
Lowry, Alice, 59, 63
Lumsden, Amy, 47
Lupyan, Gary, 53, 365, 433, 443
Lutzenberger, Hannah, 302
Lyn, Heidi, 260

Mühlenbernd, Roland, 305
Müller, Thomas, 308
Mace, Ruth, 177
MacGregor Inglis, David, 420
Machanda, Zarin, 420
MacIntosh, Andrew, 420
Mai, Anna, 374
Majolo, Bonaventura, 420
Malassis, Raphaelle, 291
Mardus, Eva, 249
Marshall, Sophie, 420
Martin, Alexander, 263
Martins, Mauricio D., 371
Martins, Pedro Tiago, 266, 417
Matsumoto, Daiki, 269
Matzinger, Theresa, 272
McGuigan, Nicola, 334
Meade, Andrew, 177
Medam, Tiphaine, 291
Meguerditchian, Adrien, 12
Mendoza Straffon, Larissa, 276
Mengede, Janine, 162, 280
Mercier, Stephanie, 420
Meyer, Antje, 359
Miaretsoa, Longondraza, 447
Michaud, Jérôme, 200, 283
Micheletta, Jérôme, 420
Moisik, Scott R., 67
Montant, Marie, 291
Montemurro, Kathryn, 296

Moran, Steven, 128
Moriano Palacios, Juan, 294
Morin, Olivier, 308
Morisseau, Tiffany, 308
Morrow, Hannah, 53
Motamedi, Yasamin, 296
Motes Rodrigo, Alba, 299
Mudd, Katie, 302
Muller, Martin, 420

Nölle, Jonas, 318, 321
Naas, Isabelle, 155
Naegeli, Danielle, 225
Nazarian, Bruno, 12
Nemeth, Dezso, 451
Nevens, Jens, 311
Nikolaev, Dmitry, 128
Notman, Hugh, 420
Noukhovitch, Michael, 314
Noveck, Ira, 308

O'Brien, Colleen, 90
O'Rourke, Thomas, 324
Oh, Yoon Mi, 44
Okanoya, Kazuo, 93, 328
Okuda, Shimpei, 168
Ordin, Mikhail, 331
Ostnera, Julia, 420
Ouattara, Karim, 420

Padden, Carol, 2, 84
Pagel, Mark, 177
Panayotov, Nikolay, 223
Papa, Aliki, 334
Parncutt, Richard, 337
Pavelka, Mary S. M., 420
Peña, Paola, 171
Peckre, Louise R., 420
Pellegrino, François, 44
Perfors, Amy, 102
Perlman, Marcus, 209
Petersdorf, Megan, 420
Pleyer, Michael, 141, 345
Poli-Nardi, Laetitia, 59
Polyanskaya, Leona, 331
Prior, Nora, 105
Pyers, Jennie, 74

Quéré, Enya, 155
Quintero, Fredy, 420

Rabagliati, Hugh, 197
Ramos-Fernández, Gabriel, 420
Rasenberg, Marlou, 356
Ravignani, Andrea, 56, 59, 63, 162
Raviv, Limor, 84, 252, 359, 440
Ren, Yi, 131
Renaud, Luc, 12
Rey, Arnaud, 451
Ridley, Amanda R., 81
Riede, Felix, 457
Robbins, Martha M., 420
Roberts, Gareth, 362
Roberts, Sean, 405, 443
Roebuck, Hettie, 365
Roettger, Timo B., 368
Rojo, Sergio Gonzalez De La Higuera, 457
Rosselló, Joana, 371
Rubio-García, Ana, 59, 63, 162

Sóskuthy, Márton, 368
Sainburg, Tim, 374
Salazar Casals, Anna, 63, 162
Saldana, Carmen, 377
Salmi, Roberta, 420
Samuel, Arthur, 331
Sandler, Wendy, 84
SanMiguel, Iria, 429
Sasahara, Kazutoshi, 168
Sato, Asha, 108

Sawallis, Thomas R., 22
Schülke, Oliver, 420
Schalz, Sabrina, 380
Schamberg, Isaac, 420
Schel, Anne, 255
Schembri, Adam, 390
Schlechtweg, Dominik, 393
Schouwstra, Marieke, 3, 24
Schulte Im Walde, Sabine, 393
Schutte, Michael, 249
Scott-Phillips, Thom, 144, 405
Segovia-Martín, José, 408
Sein, Julien, 12
Semenuks, Arturs, 477
Semple, Stuart, 420
Senghas, Ann, 84
Shcherbakova, Olena, 411
Shilton, Dor, 414
Silk, Joan B., 420
Silvente i Font, Sara, 417
Skirgård, Hedvig, 411
Slocombe, Katie, 255, 420
Smith, Andrew Marcus, 423
Smith, Kenny, 111, 131, 197, 225, 237, 321, 377, 426
Sommer, Kees, 454
Sosa-Lopéz, J. Roberto, 420
Specker, Eva, 272
Spike, Matthew, 47, 171
Sroka, Marlene, 59
Street, Sally, 420
Sturm, Stefanie, 429
Sulik, Justin, 171, 433

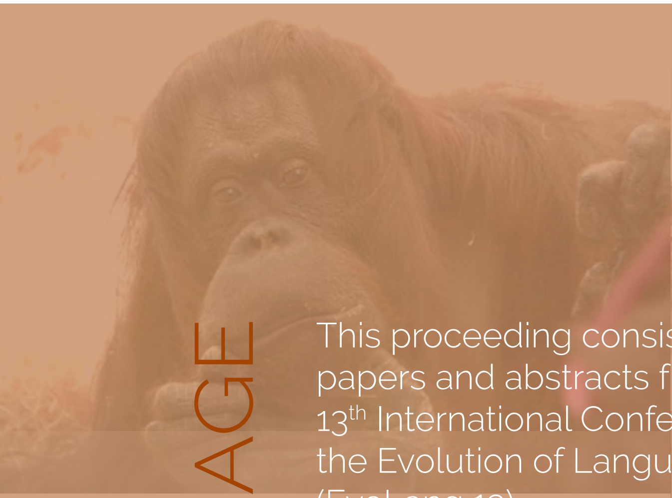
Tal, Shira, 437
Tamariz, Monica, 223, 334, 408
Tehrani, Jamie, 5
Tennie, Claudio, 299
Thompson, Bill, 56, 440, 443
Titov, Ivan, 131
Torti, Valeria, 420, 447
Tosatto, Laure, 451
Towns, Ross, 454
Townsend, Simon W., 81, 255
Tylén, Kristian, 318, 457

Valente, Daria, 420, 447
Van de Velde, Freek, 87
Van De Waal, Erica, 420
van Duijn, Max, 454
Van Eecke, Paul, 311
van Schaik, Carel P., 122
Varola, Mila, 59
Velly, Lionel, 12
Ventura, Raffaella, 420
Verhoef, Tessa, 454, 460
Vernes, Sonja C., 6, 162, 165, 249, 280
Villanueva, Stella, 162
von Heiseler, Till Nikolaus, 463

Wacewicz, Slawomir, 305
Watson, Stuart K., 81
Wedel, Andrew, 234, 471
Weyher, Anna H., 420
White, Stephanie, 93
Widmer, Paul, 181
Wiegrebe, Lutz, 249
Wiesmann, Maximilian, 165
Wilke, Claudia, 420
Winters, James, 308
Woensdregt, Marieke, 220, 474
Wrangham, Richard, 420
Wu, Ya-Ning, 477

Young, Christopher, 420

Zahn, Thiago Macek Gonçalves, 480
Zanoli, Anna, 420, 447
Zaslavsky, Noga, 484
Zhang, May, 177
Zuberbühler, Klaus, 255, 420
Zywiczynski, Przemyslaw, 305



THE EVOLUTION *of* LANGUAGE

This proceeding consists of 124 papers and abstracts from the 13th International Conference on the Evolution of Language (EvoLang 13).

As the leading international event on the evolution of language, EvoLang promotes a multidisciplinary approach to the study of language and communication systems. The contributions of this volume are written by scholars from all continents and from various fields which include biology, psychology, archaeology, genetics, neuroscience, philosophy and linguistics, among many others.