

07|08

Research Report

Max Planck Institute for Psycholinguistics





07|08

Research Report

Max Planck Institute for Psycholinguistics

Colofon

Concept Myrna Tinbergen, Intercom journalistiek & voorlichting, Nijmegen

Coordination and editing Evelyn Giering, Myrna Tinbergen, Ad Verbunt

Design Kyah ter Haar Romeny, Nijmegen

Photography Bert Beelen Fotografie, Nijmegen, MPI for Psycholinguistics

Print Drukkerij Efficiënt, Nijmegen

© Max Planck Institute for Psycholinguistics Nijmegen, April 2009

www.mpi.nl

Contents

Preface	05
Organisation of the institute	06
Honours and awards	07
PhD completions	08
The research projects of the institute	09
Project 1 <i>Decoding continuous speech</i>	10
Project 2 <i>Language in action</i>	12
Project 3 <i>The dynamics of multilingual processing</i>	14
Project 4 <i>Multimodal interaction</i>	16
Project 5 <i>Categories across language and cognition</i>	18
Project 6 <i>Information structure in language acquisition</i>	20
Project 7 <i>Unification</i>	22
Project 8 <i>Phonological learning for speech perception</i>	24
Project 9 <i>Language documentation and data mining</i>	26
Research group <i>Communication before language</i>	29
Research group <i>Adaptive listening</i>	30
Research group <i>Comparative cognitive anthropology</i>	31
Research group <i>Evolutionary processes in language and culture</i>	32
Max Planck Fellow <i>Syntax, typology and information structure</i>	33
Technical group	34
Events and activities	36

Preface

The MPI for Psycholinguistics is entering a new era. The institute's Research Report, accordingly, has been restyled to match.

There was no Annual Report 2007, as you (loyal reader) may have noticed. There will be no more Annual Reports; the Research Report is now biennial. The present issue covers 2007 and 2008. Also, the report looks rather different. This is a change mainly driven by the availability of so much information about our work on our website, nowadays. There seems little point in duplicating all this in print; instead, the report will focus on highlights and tell you where to look for more. Some things haven't changed, however; the main chapters of the report are, as before, devoted to the projects in which our research is organised. See page 9 for more about this.

And this new era, then? It is a time of change all round. The institute is expanding; and it is entering into new relationships. First, the next director has just been appointed. The last time we had a change of director was in 2006, when founding director Pim Levelt retired (in May) and Peter Hagoort began a new group (in November). Like many other Max Planck Institutes (and other institutions), we have realised that the best way to keep productivity at a maximum and to use resources to the full, however, is to overlap: a new group starts up as an old group winds down. The current Language Comprehension group will close early in 2013. But a new group, directed by professor Antje Meyer, currently of the University of Birmingham, will begin at the end of 2009.

But that is not all. We have also received the green light to begin a further department, in the area of language and genetics. The Max Planck Society has appointed a committee to select a director for this new department, and the selection process will begin during 2009 too. This expansion will necessitate extension of our current building, of course; that too has already been approved. Exciting times ahead!

The Max Planck Society has a system of Max Planck Research Groups, which allow young scientists of exceptional promise to set up an independent research programme for a period of several years. In the first 27 years of its existence, the MPI for Psycholinguistics had one of these groups. In 2008, we acquired four more! They report here for the first time (pp. 29 to 32). Also we acquired our first Max Planck Fellow; this is a senior scientist based at a university who heads a collaborative project carried out by a small group at the institute. Professor Robert van Valin, of the University of Düsseldorf, who heads this group, also presents his first report in this issue (p. 33). The addition of these groups made extra temporary accommodation necessary; we now have two prefabricated container buildings next to our main building for the time being. (The little one that arrived in 2003 was dubbed the Mini-Planck. The slightly larger one which was set between it and the main building has acquired the name Midi-Planck...of course.)

In the Radboud University Nijmegen, the graduate schools in the areas of psychology and neuroscience have merged in the new Donders Institute for Brain, Cognition and Behaviour. The MPI for Psycholinguistics is a partner in this new institute. The Donders Institute, with the MPI, will also participate in a new Max Planck International Research School, which was also approved in 2008; it will start up in 2009.

As if this weren't enough excitement, now read what's in the pages to come!

Anne Cutler
Managing director 2009-2010
April 2009



The current board of directors (f.l.t.r.): Wolfgang Klein, Stephen Levinson, Anne Cutler and Peter Hagoort

Organisation of the institute 2007-2008



Honours and awards

Directors

Stephen Levinson (*managing director*)
 Anne Cutler
 Peter Hagoort
 Wolfgang Klein

Director emeritus

Willem Levelt

Max Planck Research Groups

Michael Dunn (*head*)
 Daniel Haun (*head*)
 Ulf Liszkowski (*head*)
 Andrea Weber (*head*)

Max Planck Fellow

Robert van Valin (*Heinrich Heine U. Düsseldorf*)

External scientific members

Manfred Bierwisch
 Pieter Muysken

Scientific council: 2007

Dan Slobin, chair (*U. California at Berkeley*)
 Kathryn Bock (*U. Illinois*)
 Herbert Clark (*Stanford U.*)
 Nick Evans (*U. Melbourne*)
 Kenneth Forster (*U. Arizona*)
 Beatrice de Gelder (*U. Brabant, Tilburg*)
 Edward de Haan (*Utrecht U.*)
 Hans Kamp (*U. Stuttgart*)
 John Lucy (*U. Chicago*)
 Joanne Miller (*Northeastern U., Boston*)
 John Schumann (*U. California, Los Angeles*)

Scientific council: 2008

Dan Slobin, chair (*U. California at Berkeley*)
 David Birdsong (*U. Texas at Austin*)
 Herbert Clark (*Stanford U.*)
 Carol Fowler (*Yale U.*)
 Dedre Gentner (*Northwestern U.*)
 Edward de Haan (*Utrecht U.*)
 Aafke Hulk (*U. Amsterdam*)
 Manfred Krifka (*Humboldt U. Berlin*)
 Robert Ladd (*U. Edinburgh*)
 Thomas Münte (*U. Magdeburg*)
 Eve Sweetser (*U. California at Berkeley*)

Head of technical group

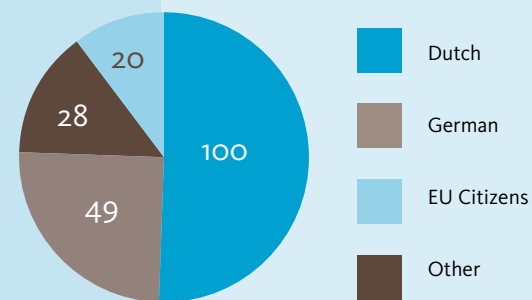
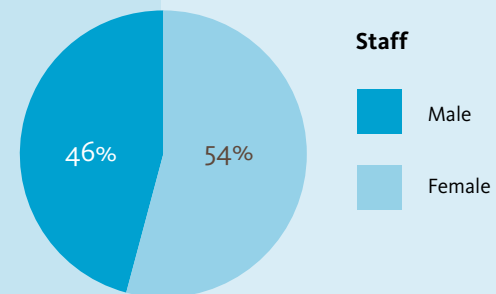
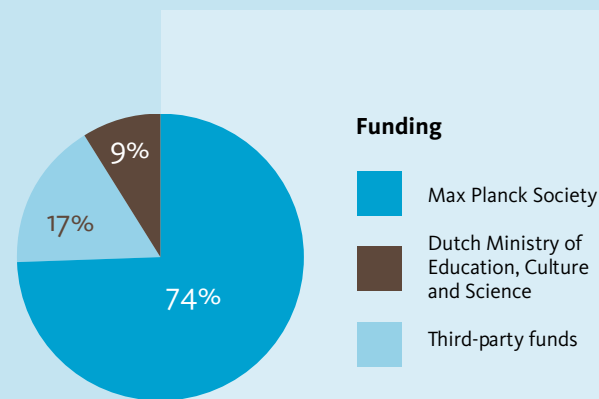
Peter Wittenburg

Head of administration

Paul Lommen

Head of library

Karin Kastens



2007

Melissa Bowerman was feted with a symposium in her honour and presented with a 'Festschrift' at the Boston University Conference on Language Development.

Anne Cutler was elected Foreign Member of the American Philosophical Society.

Ian FitzPatrick was elected MPG PhD net representative for Humanities.

Peter Hagoort was awarded an honorary degree in science for contributions to the cognitive neuroscience of language by the University of Glasgow.

Laura Herbst won the Peter Ladefoged prize for the best paper by a student participant at the XVth International Congress of Phonetic Sciences.

Laura Herbst received a Marie Curie Fellowship for early stage researchers with the Edinburgh Speech Science and Technology programme.

Alexandra Jesse received the Innovational Research Incentives Scheme VENI grant of the Netherlands Organisation for Scientific Research (NWO), hosted by the MPI.

Kristin Lemhöfer received the Innovational Research Incentives Scheme VENI grant of the Netherlands Organisation for Scientific Research (NWO), hosted by the MPI.

Sarah Schimke received a 'Kurzstipendium für Doktoranden' of the German Academic Exchange Service (DAAD).

2008

Susanne Brouwer won the poster prize at Laboratory Phonology 11 in Wellington.

Susanne Brouwer was awarded a grant from the Hugh Knowles Leadership Fund to do research at Northwestern University in Evanston (USA) for 6 months.

Anne Cutler was elected Foreign Associate of the National Academy of Sciences (US).

Anne Cutler was elected Honorary Fellow of the Australian Academy of the Humanities.

Anne Cutler was Kanizsa Lecturer at the University of Trieste (Italy).

Mirjam Ernestus became member of the Young Academy of the Royal Netherlands Academy of Arts and Sciences (KNAW).

Rik van Gijn received the Innovational Research Incentives Scheme VENI grant of the Netherlands Organisation for Scientific Research (NWO), hosted by the Radboud University Nijmegen.

Peter Hagoort was awarded the Senior Heymans Prize of the NIP (the Dutch Professional Association of Psychologists).

Peter Indefrey was awarded a Heisenberg Professorship of the German Research Foundation (DFG).

Wolfgang Klein was elected member of the Deutsche Akademie für Sprache und Dichtung in Darmstadt.



PhD completions



Heidrun Bien: 'On the production of morphologically complex words with special attention to effects of frequency'.

Amanda Brown: 'Crosslinguistic influence in first and second languages: Convergence in speech and gesture'.

Jidong Chen: 'The acquisition of verb compounding in Mandarin Chinese'.

Daniel Haun: 'Cognitive cladistics and the relativity of spatial cognition'.

Suzanne van der Feest: 'Building a Phonological Lexicon: The acquisition of the Dutch voicing contrast in perception and production'.

Martijn Goudbeek: 'The acquisition of auditory categories'.

Valesca Kooijman: 'Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence'.

Victor Kuperman: 'Lexical processing of morphologically complex words: An information-theoretical perspective'.

Pamela Perniss: 'Space and iconicity in German Sign Language' (DGS).

Mark Pluymaekers: 'Affix reduction in spoken Dutch'

Anita Wagner: 'Phoneme inventories and patterns of speech sound perception'.

Claudia Wegener: 'A grammar of Savosavo: A Papuan language of the Solomon Islands'.

The research projects of the institute

The discipline of psycholinguistics has its roots in two research fields: Psychology and linguistics. Traditionally, psychologists work in faculties of social sciences, linguists in humanities faculties; our institute is one of the very few places in the world where psycholinguists from varying backgrounds work together in the same lab, the same corridor, the same project. The institute's research is organised so as to capitalise on the rare opportunity this presents: There are a small number of overarching projects each of which addresses a single general theme. Projects run for several years, and the PhD projects at the institute are embedded in a particular project; each project has a number of subprojects extending the general theme, and each project draws on both psychological and linguistic expertise (not to mention anthropological, neuroscientific, computational, biological, genetic or medical expertise, if appropriate!).

The nine institute projects embracing the research conducted in 2007 and 2008 are described in summary reports in the following pages; following them, there are then summary reports from the independent research projects of the Max Planck Research Groups and the Max Planck Fellow. Much more information about each project, and in particular a full list of publications, may be found on our website: www.mpi.nl.



Project 1

Decoding continuous speech

Project coordinator James McQueen

Project members Attila Andics, Bettina Braun, Susanne Brouwer, Anne Cutler, Falk Huettig, Alexandra Jesse, Claudia Kuzla, James McQueen, Holger Mitterer, Dennis Norris (MRC Cognition and Brain Sciences Unit, Cambridge), Eva Reinisch, Matthias Sjerps, Lara Tagliapietra, Marco van der Ven



Goals of the project

To understand talkers' messages, listeners must crack the speech code. Behind the apparent effortless of speech recognition (we seem simply to hear words) lurks a complex problem. Spoken words vary enormously in their realisation, for example because of differences among talkers, speaking rates, and sentential contexts. Words must also be segmented out of a continuous acoustic signal. The goal of this project, which ended in 2008, was thus to understand how listeners decode speech. Methods included behavioural and neuroscientific experimentation, cross-linguistic comparisons, and computational modelling.

Subproject: Printed-word eye-tracking

The project pioneered the development of a new eye-tracking methodology, using visual displays containing printed words rather than pictures (see Figure 1). The timing and proportion of fixations to pictures or words reveal which lexical hypotheses the listener considers as speech unfolds over time. But experiments with pictures have the unfortunate constraint that the spoken materials must refer to picturable objects. More stimuli can be tested using printed-word displays. Initial work established the validity of the new paradigm (see Annual Report 2005). The project has since used printed-word eye-tracking to answer several questions.

(1) *How are reduced forms recognised?* In her PhD work, Susanne Brouwer is studying the recognition of reduced forms (e.g., *puter* instead of the canonical form *computer*). Listeners heard sentences while looking at displays including a target word mentioned in the current sentence (e.g., *computer*), a competitor phonologically similar to the canonical form (e.g., *companion*), and a competitor similar to the reduced form (e.g., *pupil*).

Listeners' fixations to the two types of competitor varied depending on speech style (e.g., laboratory vs. spontaneous speech) and target pronunciation (e.g., *computer* vs. *puter*). These results suggest that listeners recognise reduced forms in part by varying, across listening situations, the set of competitors that they actively consider.

(2) *When is lexical stress information used?* Eva Reinisch's PhD project has investigated when Dutch listeners use suprasegmental lexical stress information to recognise words. Fixations on printed targets such as *OCtopus* (octopus, capitals indicate stress) were more frequent than fixations on segmentally overlapping but differently stressed competitors (*okTOber*, October). Critically, this occurred before segmental information could disambiguate the words. Furthermore, initially stressed words were stronger competitors than non-initially stressed words prior to segmental disambiguation. Listeners appear to recognise words optimally by immediately using all relevant information.

(3) *How is phonological knowledge stored in the lexicon?* Holger Mitterer tested the claim that lexical representations are underspecified with respect to default phonological features (e.g., that the [t] of *tin* is not specified lexically as coronal and voiceless). Contrary to this claim, (a) listeners did not look more at *tin* when they heard *pin* than they looked at *pin* when they heard *tin*, and (b) they looked more at *tin* when they heard *pin* than when they heard *bin*. These results suggest instead that the phonological forms of words are fully specified in the lexicon.

(4) *Is retrieval of lexical-conceptual knowledge automatic?* Falk Huettig and James McQueen found that participants looked at pictures of phonological, semantic, and visual-feature competitors of spoken words (e.g., looks to a beaver, a fork and a bobbin as *beaker* was being heard). When the pictures were replaced with their printed names, there were looks only to the phonological competitors. When these were removed from the display, there were looks to the semantic but not the visual-feature competitors. Retrieval of components of lexical knowledge is thus not automatic, but instead depends on the situation listeners find themselves in.

Subproject: No feedback during speech perception

James McQueen, Alexandra Jesse and Dennis Norris found that lexical knowledge (e.g., that the ambiguous fricative

ending *Christma?* should be [s]) did not influence perception of stops following those fricatives. A previous apparent demonstration of this kind of lexical influence was shown to depend on whether practice stimuli were words or nonwords. When both were used, there was no effect. Across experiments, however, there were always lexical effects on fricatives (e.g., the final sound of *Christma?* was identified as [s]). The prelexical level of speech processing appears to be sensitive to practice-induced phoneme-sequence biases. Critically, the within-trial dissociation (lexical involvement in the fricative but not the stop decisions) challenges interactive models in which feedback to the prelexical level should cause both effects. As Shortlist B (see below) predicts, there thus appears to be no feedback during speech perception.

Other subprojects

- Attila Andics' PhD project examines voice category learning. Representations of new voices are flexible, and involve segmental and nonsegmental talker-specific details; fMRI experiments reveal that this flexibility is subserved by voice-selective neural populations.
- In his PhD work, Matthias Sjerps has tested whether the adjustment process responsible for vowel normalisation operates only on speech. Adjustments are made to nonspeech stimuli, but only if those stimuli have spectrotemporal similarity to speech.
- James McQueen and Dennis Norris completed the development of Short-

list B, a Bayesian model. Shortlist B's success in simulating key findings suggests that listeners make optimal Bayesian decisions during word-recognition.

Selected publications

- Cho, T., McQueen, J.M., & Cox, E.A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, 35(2), 210–243.
- Huettig, F., & McQueen, J.M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57(4), 460–482.
- Kuzla, C., Cho, T., & Ernestus, M. (2007). Prosodic strengthening of German fricatives in duration and assimilatory devoicing. *Journal of Phonetics*, 35(3), 301–320.
- Mitterer, H., Yoneyama, K., & Ernestus, M. (2008). How we hear what is hardly there: Mechanisms underlying compensation for /t/-reduction in speech comprehension. *Journal of Memory and Language*, 59(1), 133–152.
- Norris, D., & McQueen, J.M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, 115, 357–395.

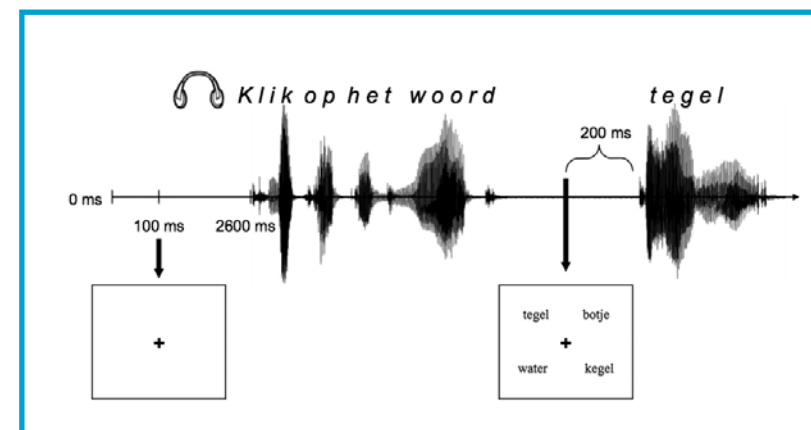


Figure 1 Printed-word eye-tracking. Listeners fixate a cross, then hear a sentence (e.g., *Klik op het woord 'tegel'*, click on the word 'tile'). A display of four printed words (here including 'tegel' and the phonological competitor 'kegel' 'skittle') appears before target onset. Fixations over time to these displays are recorded using a head-mounted camera (see photos).



Project 2

Language in action

Project coordinator Jos van Berkum

Project members Petra van Alphen, Jos van Berkum, Dannie van den Brink, Daniel Casasanto, Laura Staum-Casasanto, Dieuwke de Goede, Peter Hagoort, Jana Hanulova, Laura Menenti, Asli Özyürek, Cathelijne Tesink, Roel Willems



Goals of the project

Language helps people interact with their social and physical environment – it is for doing things. The Language in action project investigates the basic neural and cognitive architecture of language comprehension and production, while taking this richer context into account. The project takes an explicit cognitive neuroscience stance, and data are mainly acquired through well-controlled experimentation in the laboratory. Current research focuses on the neural architecture of pragmatic inferencing (indirect replies; establishing reference; discourse – world knowledge interaction), on the language-valence interface (interaction with value-systems and mood) and on language-body interactions (how bodily experience shapes language processing).

Subproject: Morally objectionable language

People disagree over things of fundamental importance, such as whether euthanasia is acceptable, whether it is okay to joke about somebody's religion, and whether one's country should shut out economic refugees. The moral values behind these disagreements are frequently debated with language, and – in attitude surveys – probed through language. In this subproject we use EEG and MEG to study how rapidly a person's values are brought to bear on processing as they read an attitude survey statement, what the neural consequences of such value-based processing are, and whether values can in fact influence the ongoing linguistic-semantic analysis. In one study, we recorded EEG while respondents from contrasting political-ethical backgrounds completed a realistic attitude survey on drugs, medical ethics, social conduct and other issues. Our results show that value-based disagreement is unlocked by language extremely rapidly, within 200 – 250 ms after the first word at which a statement begins

to clash with the reader's value system (see Figure 1). Strong disagreement also elicits the standard ERP effect to negatively valenced stimuli (LPP effect), and, most interestingly, influences the ongoing analysis of meaning (N400 effect), indicating that even very early processes in language comprehension are sensitive to a person's value system. Our results

testify to rapid reciprocal links between neural systems for language and for valuation, and are consistent with the idea that valence is an integral part of early language interpretation. They also show that a real-world arena of language use is amenable to cognitive neuroscience research.

Subproject 2: The body-specificity hypothesis

How might bodily experience shape thinking and language processing? According to the body-specificity hypothesis, if thinking and language processing are 'embodied', then people with different kinds of bodies, who interact with their physical environments in different ways, should form correspondingly different concepts and word meanings. One set of experiments investigated the neurocognitive representation of

action verb meanings in right- and left-handers. Embodied theories posit that understanding a verb like *throw* involves unconsciously simulating the action *throwing*, in regions of the brain that support motor planning. We used fMRI to compare premotor activity correlated with verb understanding in right- and left-handers (see Figure 2). Right-handers preferentially activated left premotor cortex when reading manual action verbs (compared with non-manual action verbs), whereas left-handers preferentially activated right premotor cortex. This finding helps refine theories of embodied semantics, suggesting that unconscious mental simulation during language processing is body-specific: Right- and left-handers, who perform actions differently, use correspondingly different areas of the brain for representing action verb meanings.

Another set of experiments investigated relationships between the emotional content of speech and the hand used to make spontaneous co-speech gestures. The US presidential debates from the past two elections provided a test bed, as both candidates from 2004 were right-handers (Kerry, Bush), and both candidates from 2008 were left-handers (Obama, McCain). Right-handers produced a greater proportion of right-hand gestures during clauses with positive emotional valence, and left-hand gestures during clauses with negative valence; left-handers showed the opposite pattern, consistent with the body-specificity hypothesis.

Other subprojects

- Pragmatic inferencing in the comprehension of indirect replies.
- The neural substrate for establishing reference.
- The interaction between discourse context and world knowledge.
- How mood affects heuristic and algorithmic aspects of comprehension.
- Metaphors in language, mind, and brain.
- How language creates 'embodied' representations of things we have never experienced directly.
- Event description in gesturing and other modalities.
- Discourse-based anticipation of upcoming language.

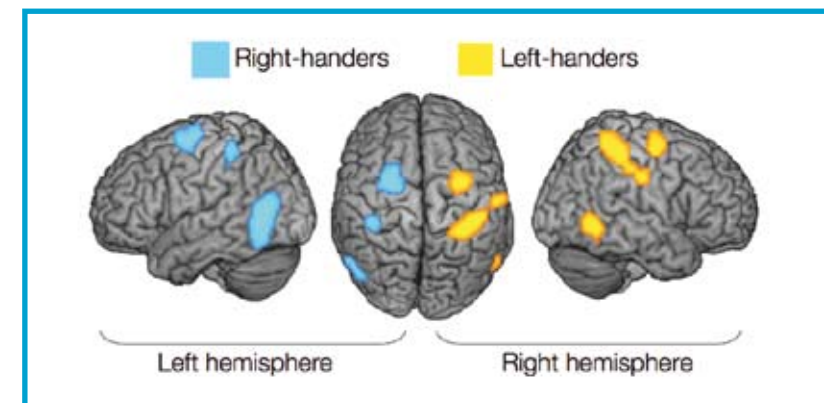


Figure 2 fMRI activation showing that right- and left-handers use different cerebral hemispheres when imagining the actions named by manual action verbs like 'grijpen' ('grasp') or 'gooien' ('throw').

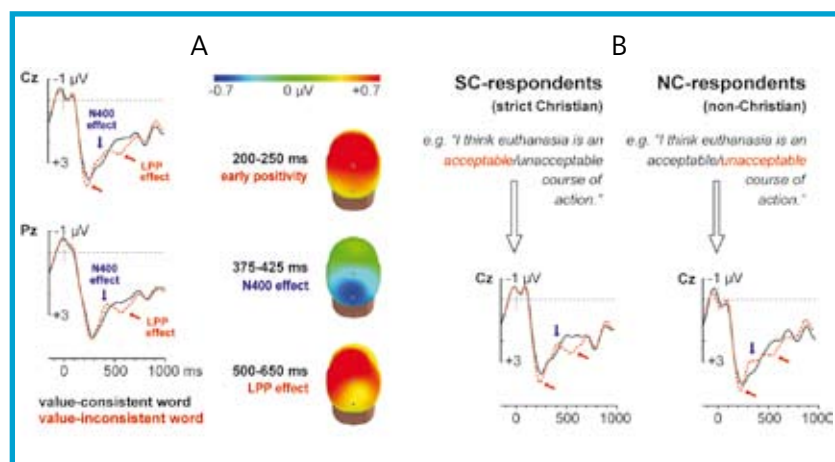


Figure 1 The EEG brain signature of strong disagreement during reading of attitude survey statements (A) pooled across two groups of respondents, with scalp distributions for differential effects, and (B) per group.

Selected publications

- Hagoort, P., & van Berkum, J.J.A. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society, B*, 362, 801–811.
- van Berkum, J.J.A. (2008). Understanding sentence in context: What brain waves can tell us. *Current Directions in Psychological Science*, 17(3), 177–182.
- van Berkum, J.J.A., van den Brink, D., Tesink, C.M.J.Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, 20(4), 580–591.
- Goldin-Meadow, S., So, W., Özyürek, A., & Mylander, C. (2008). The natural order of events: How do speakers of different languages represent events nonverbally? *Proceedings of the National Academy of Sciences*, 105(27), 9163–9168.
- Tesink, C.M.J.Y., Petersson, K.M., van Berkum, J.J.A., van den Brink, D., Buitelaar, J.K., & Hagoort, P. (2008). Unification of speaker and meaning in language comprehension: An fMRI study. *Journal of Cognitive Neuroscience*. Published online 2008-11-18.



Project 3

The dynamics of multilingual processing

Project coordinators Marianne Gullberg and Peter Indefrey

Project members Doug Davidson, Ian Fitzpatrick, Marianne Gullberg, Jana Hanulová, Frauke Hellwig, Peter Indefrey, Wolfgang Klein, Nan van de Merendonk, Georgina Oliver, Leah Roberts, Hülya Sahin, Sabine Schneider, Kim Veroude, Kirsten Weber



Goals of the project

The project studies the processing of second and third languages (L2+) during first contact, during acquisition, and at the level of high proficiency or functional bilingualism. These aspects of L2+ processing are explored by use of different methodologies, including reaction time experiments, eye-tracking and gesture analysis. Experimental work on L2+ processing using neuroimaging techniques (fMRI, ERP) is conducted at the Donders Institute for Brain, Cognition and Behaviour in Nijmegen.

Subproject 1: Changes in brain responses during morphosyntactic learning

This subproject investigates how quickly the learning of a morphosyntactic rule of a second language can result in changes of an electrophysiological brain response to syntactic violations (P600). In Dutch native speakers learning German adjective declension we found a rapid development of a P600 response to declension class violations in a paradigm where learners were presented with phrases like *mit dem kleinen/*kleinem Kind* [with the small child] and were asked to judge their correctness. In a pretest, incorrect adjective forms redundantly marking case, gender, and number elicited no violation response. After instruction of the rule, participants received feedback on their judgements in a training session and were finally tested again one week later. Already in the training session the brain responses of Dutch learners began to look like those of German native speakers (see Figure 1). We also found that the magnitude of a brain response

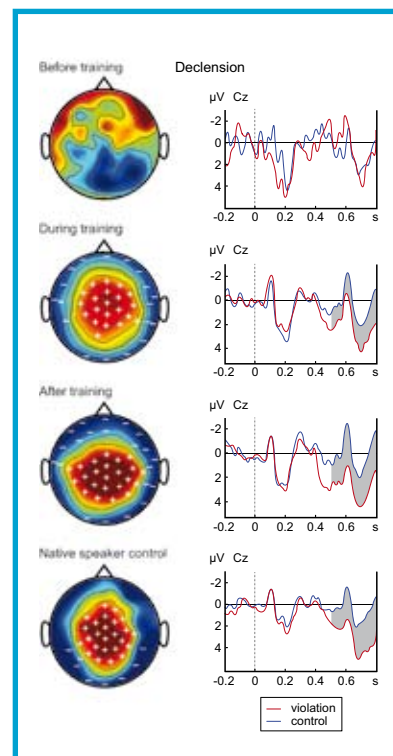


Figure 1 Left: topographies for the declension violation response (violation-control) in the P600 time window. Electrodes within significant clusters are plotted for both positive (+) and negative (-) average potentials. Right: ERP traces (channel Cz) from the onset of the critical adjective (blue curve: correct; red curve: declension violation). Significant P600 amplitude differences are shaded in gray.

to errors followed by negative feedback (error-related negativity) in the training session predicted the learning success in the test session.

Subproject 2: First language activation during second language sentence listening

It is known that words from a bilingual speaker's two languages interact in the mental lexicon. Activation of words from the 'wrong' language might interfere with comprehension. Research in this subproject investigates under what conditions and to what extent words from the first language are activated when listening to sentences in the second language. Using an electrophysiological measure of semantic integration into the sentence context (N400), we found that semantically inappropriate L2 words that sound initially like an appropriate L2 word lead to a delay in the N400 response indicating semantic incongruity, but L2 words that sound initially like an appropriate L1 word do not (see Figure 2). These findings show that semantic integration in non-native listening can start on the basis of word initial phonemes, but that spuriously elicited L1 lexical candidates are not available for semantic integration in L2 speech comprehension. However, results differ when not only initial phonemes but the whole word may be from the first language. Another ERP study showed that the Dutch meaning

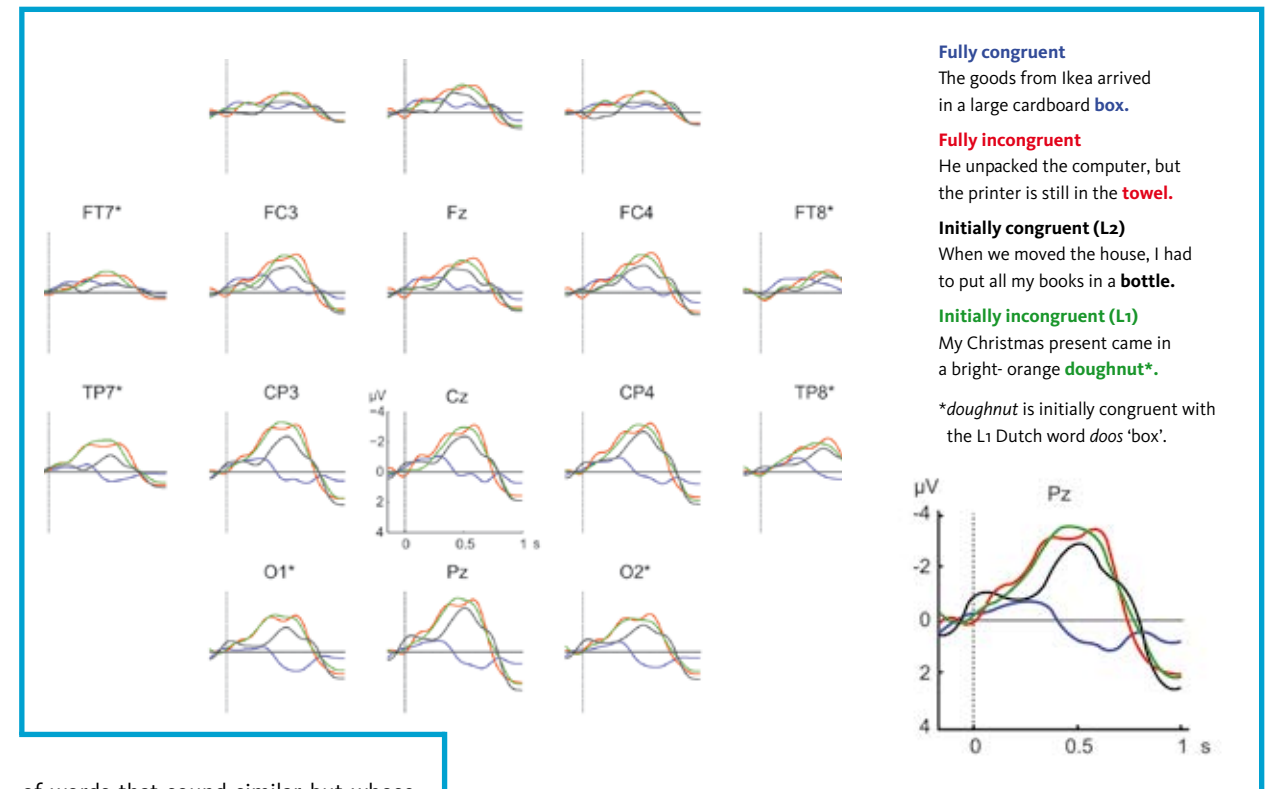


Figure 2 N400 responses to critical words shown in coloured frames for example sentences of the four conditions. Grand average waveforms from the onset of critical words on 16 scalp electrodes. N400 onset was delayed for critical words with initial phonemic overlap with a fully congruent sentence-final L2 word (black curve).

of words that sound similar but whose meanings differ in the two languages (e.g., *pet*, meaning 'cap' in Dutch) is activated when listening to English sentences. This holds even when the sentence context favours the English meaning and is incompatible with the Dutch meaning, as in: His cat was his favourite pet.

Other subprojects

- The effects of minimal exposure to an unknown L2.
- The short-term learning of Dutch crossed dependencies.
- The processing of aspect in L2.
- The processing of L2 word order.
- The processing of null and overt pronouns in L2 discourse.
- The restructuring of semantic representations in L2 production.
- The time course of processing stages in L2 word production.
- Code-switching in production.

Selected publications

- Davidson, D.J., & Indefrey, P. (2007). An inverse relation between event-related and time-frequency violation responses in sentence processing. *Brain Research*, 1158, 81–92.
- Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of Manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, 30(2), 225–251.
- Davidson, D.J., Indefrey, P., & Gullberg, M. (2008). Words that second language learners are likely to hear, read, and use. *Bilingualism: Language and Cognition*, 11(1), 133–146.
- Roberts, L., Gullberg, M., & Indefrey, P. (2008). On-line pronoun resolution in L2 discourse: L1 influence and general learner effect. *Studies in Second Language Acquisition*, 30(3), 333–357.
- Indefrey, P., & Gullberg, M. (eds.) (2009). *Time to speak: The cognitive and neural prerequisites for time in language*. Malden, MA: Wiley-Blackwell.



Project 4

Multimodal interaction

Project coordinators Nick Enfield and Jan Peter de Ruiter

Project members Penelope Brown, Nick Enfield, Sonja Gipper, Kaoru Hayano, Gertie Hoymann, Olivier Le Guen, Stephen Levinson, Lilla Magyari, Federico Rossano, Jan Peter de Ruiter, Mark Sicoli, Tanya Stivers



Goals of the project

This project investigates language in its primordial context – face-to-face conversational interaction – the context in which language is learnt, and predominantly used. Work in the project is focused on the idea that there are strong constraints and special faculties underlying interactional uses of language, including principles of mutual orientation, coordination, turn-taking, information tracking and timing. Current work is mostly based on cross-linguistic and cross-cultural comparison of interactive language use to extract universal patterns in domains where cultural variation may be greatest. The following illustrates the kind of research involved.

Subproject: Questions

Extensive conversational data have been collected from 10 languages in five continents, and from each of these languages 350 questions and their responses have been extracted. It was found that in most languages, Yes-No questions were the most common question type. In only one of the languages – !Akhoe Haillom, a language of Namibia – were content questions (questions about 'what', 'where', 'why', 'who', etc.) more common.

Questions are coded in different ways in the 10 languages. In all languages, some of the questions have no formal lexico-morpho-syntactic marking at all, but the frequency of this varies substantially. With respect to Yes-No questions, speakers of !Akhoe Haillom and Korean used non-marked question forms less than 2% of the time, whereas speakers of Japanese and Danish did so more than 20% of the time and speakers of Yélfí-Dnye did so in a majority of cases – 65% of the time. This means that a substantial proportion of questions across all languages are recognised as such by general inference. Often,

this occurs when what looks like a plain statement about the addressee's domain of knowledge is understood as a question (e.g., 'You have a meeting at three o'clock').

Questions are typically associated with the pragmatic action of requesting information (e.g., 'How would I get to the theatre from here?'). However, across most languages speakers use questions to

implement a range of other social actions as well. Strikingly, requesting confirmation is the most common type of question in 4/10 languages and was the most common question function, overall accounting for 43% of all questions asked. This is notable because confirmation requests position the asking party as already thinking that something is the case (e.g., 'You're coming, aren't you?'). However, in !Akhoe Haillom confirmation requests were absent entirely.

A third action that questions were commonly used for across languages was the initiation of 'repair' of a problem in speaking or understanding (e.g., 'What?' or 'He did what?'). Along with requests for information and confirmation, initiations of repair formed the third most common pragmatic action implemented with questions across the languages. However, in Japanese and Lao there were fewer initiations of repair than in other languages, and questions were used more often for doing

assessments (e.g., 'Isn't she a beautiful girl?' where the speaker is not actually asking a question, but rather meaning to convey that she thinks the girl is beautiful).

Responses to questions

There is also considerable variation in the types of response that a question receives. Speakers of nearly all languages show a propensity for giving interjection responses to Yes-No questions. This means that they are more likely to answer a question like 'Is John at home?' with an interjection like 'Yes' or 'No' than with a repetition or partial repetition like 'He is' or 'He is at home'. However, in two cases (Tzeltal and !Akhoe Haillom) speakers relied heavily on repetitions (e.g., 'I am going' or 'I am') to answer questions. And in several other languages the frequency of use of repetitions in responding is higher than in European languages such as English (e.g., Japanese, Lao and Yélfí-Dnye).

The research carried out in this project is mostly done with the use of video recordings. This allows us to examine visual aspects of communicative beha-

viour such as eye gaze and hand gesture. In the Questions project, it was found that speakers of all languages rely on visible responses such as head nods, head shakes and shoulder shrugs. The study showed that speakers of certain languages are more likely than others to rely on these responses. For example, speakers of Japanese and Yélfí-Dnye are more likely to include a visible response than the other languages. Conversely, speakers of Tzeltal are less likely than speakers of other languages to include a visible response. In earlier project work, these tendencies for visual behaviour have been associated with general cultural differences in conversational style. A final aspect of the responses to questions that was measured is the timing of responses, that is, how much time passes between the end of a question and the beginning of the response. We found that variation in speed of response is a function of a range of factors, including the pragmatic action being implemented by a question, the type of response being given to the question, and speaker gaze. These accounts for variation

in speed of response held in much the same way in all languages.

The parallels across the languages can best be explained in terms of norms of timely uptake in conversation (neat transition from one speaker to the next, without significant delay or overlap), biased towards cooperative, socially aligning responses. These results suggest that conversational interaction has a largely universal substrate, which may be traced to the general bias towards cooperation in our species, a bias that underlies the very possibility of language and culture.

Other subprojects

- The ability to communicate outside a conventional code, and a possible neural substrate for this ability.
- The ability to 'project' the ending of another speaker's turn at speaking, and a possible neural substrate for this ability.
- The role played by knowledge and authority differences in the structuring of conversation.
- The role of eye gaze in the coordination of turns and sequences of conversation.
- The cultural elaboration of the human propensity for intention-attribution, in ritual and communication of a Central American culture.
- Systems of modality and evidentiality, as encoded in the grammar of a South American language.
- Gestural communication in pre-linguistic infancy: cross-cultural perspectives.
- Prosody in question-asking.



Selected publications

Enfield, N.J., & Stivers, T. (eds.) (2007). *Person Reference in Interaction: Linguistic, Cultural, and Social Perspectives*. Cambridge: Cambridge University Press.

de Ruiter, J.P. (2007). Postcards from the mind: the relationship between speech, imagistic gesture, and thought. *Gesture* 7(1), 21–38.

Stivers, T. (2008). Stance, alignment, and affiliation during storytelling: When nodding is a token of affiliation. *Research on Language and Social Interaction*, 41(1), 31–57.

Rossano, F., Brown, P., & Levinson, S.C. (2009). Gaze, questioning, and culture. In J. Sidnell (ed.), *Comparative studies in conversation analysis*. Cambridge: Cambridge University Press.



Project 5

Categories across language and cognition

Project coordinators Asifa Majid and Gunter Senft

Project members Melissa Bowerman, Penelope Brown, Niclas Burenhult, Mark Dingemanse, Nick Enfield, Sebastian Fedden, Rik van Gijn, Peter Hagoort, Clair Hill, Falk Huettig, Olivier Le Guen, Stephen Levinson, Asifa Majid, Özge Öztürk, Disa Sauter, Gunter Senft, Shakila Shayan, Mark Sicoli, Hilario de Sousa, Sylvia Tufvesson



Goals of the project

Language has to interact with a large range of other representational systems – visual, auditory, haptic, olfactory – if we are to talk about the world. What kinds of transformations of concepts are involved in describing what we see and otherwise experience? This project is concerned with semantic categories and how they are expressed in words and grammar, as well as how variable linguistic categories hook up to other perceptual and conceptual representation systems which are phylogenetically older, ontogenetically antecedent and biologically ‘fixed’ to a much greater extent. Project members conduct cross-linguistic and cross-cultural comparison to examine the language-cognition interface.

Subproject 1: ‘Ineffables’

Why is it that language is good at describing certain states of affairs (e.g., how to get to the cinema, or the kinship relation between me and my grandfather), but very limited in others (e.g., describing smells or faces)? Ineffability – the difficulty or impossibility of putting certain experiences into words – is a topic that has been relatively neglected in the linguistic and cognitive sciences. But limits on the ability to express in words sensory and emotional experiences can tell us important things about how the mind works, how different modalities do or do not talk to one another, and how language does, or does not, interact with other mental faculties.

For example, English speakers find it extremely difficult to describe smells so that they can be uniquely identified – try to describe how your grandmother’s house smells... Or describe the taste of a banana so that someone who has never tried it could conjure up the taste sensation. Here we find ourselves at the limit of the descriptive power of language. This subproject examines the linguistic

codability of sensory and emotional experiences and seeks to understand why there might be differential ineffability of the senses. To what extent is differential ineffability a result of our cognitive architecture, a question of how language hooks up to other representational systems, and to what extent is it a question

of local cultural preoccupation such that we could have developed rich resources for talking about smells and tastes?

To address this question, a standardised set of stimuli of colour patches, geometric shapes, simple sounds, tactile textures, smells and tastes has been used to elicit descriptions from speakers of more than a dozen languages. The languages are typologically, genetically and geographically diverse, representing a wide range of subsistence modes (hunter-gatherer to industrial), ecological zones (rainforest jungle to desert) and dwelling types (rural and urban). We examine how codable the different sensory modalities are by comparing how consistent speakers are in how they describe the materials in each modality. Preliminary results suggest that differential codability may be partly the result



of cultural preoccupation, with some cultures being more visually oriented and others more olfactorily oriented.

Qualitative analyses of the data also reveal a special class of words – expressives – used to convey a vivid impression of certain sensations or sensory perceptions. These are found abundantly in Asian and African languages, as well as in some South American languages but are rare in Indo-European languages, and thus provide a special focus for this project. Straddling boundaries that have long been considered self-evident in Western thought (such as perception vs. emotion, or the traditional five-senses model), expressives provide a unique view of cultural meaning systems relating to perception and sensation. Due to their sound-symbolic and typically multimodal nature, expressives provide crucial insights about iconicity and modality to questions of (in)effability of categories across language and cognition.

Subproject 2: The interface between language and cognition

This project examines the role of language in our cognitive system. For example, previous research has shown that world knowledge affects our perceptual decisions – a yellow-orange hue on a banana is categorised as being more yellow than the same hue on a carrot (see Phonological learning for speech perception project). But the mechanism by which

this occurs is unknown. Does a banana appear more yellow than it actually is because we know (declaratively) that ‘bananas are yellow’, or do we assume that the hue on this banana is probably similar to the hues we (perceptually) remember from encounters with other bananas? Declarative and perceptual memories are perfectly correlated under usual circumstances. But traffic lights in Germany and the Netherlands provide a test bed for teasing these two factors apart. The topmost and bottommost colours of traffic lights are described with cognate colour terms for ‘green’ and ‘red’ in the two languages, but the middle light is subject to variation: Germans use their equivalent of ‘yellow’ (i.e., *gelb*), but Dutch speakers use the term for their national colour ‘orange’ (*oranje*). This diversity arises even though the perceptual memories of Dutch and Germans must be similar, because traf-

fic lights in the European Union are bound by a European norm. When asked to categorise a yellow-orange hue continuum on the middle traffic light, German speakers gave relatively more yellow responses than Dutch speakers, showing that declarative – and not perceptual – memory underlies how we categorise colours on objects. This demonstrates the powerful role language plays in how we categorise everyday objects.

Other subprojects

- Comparative codability of emotions.
- The cross-cultural development of sensory and emotional categories.
- The role of grammatical categories in object categorisation and attention.
- Event categorisation in language and cognition.
- Spatial cognition and spatial language across cultures.

Selected publications

Ameka, F., & Levinson, S.C. (eds.) (2007). The typology and semantics of locative predicates: Posturals, positionals and other beasts. Positionals: The state of the art. *Linguistics* [Special issue] 45(5/6), 847–1151.

Majid, A., & Bowerman, M. (eds.) (2007). Cutting and breaking events: A cross-linguistic perspective. *Cognitive Linguistics* [Special issue] 18, 133–337

Burenhult, N. (ed.) (2008). Language and landscape: Geographical ontology in cross-linguistic perspective. *Language Sciences* [Special Issue] 30(2/3), 135–382.

Majid, A., Boster, J. S., & Bowerman, M. (2008). The cross-linguistic categorisation of everyday events: A study of cutting and breaking. *Cognition*, 109(2), 235–250.

Senft, G. (ed.) (2008). *Serial Verb Constructions in Austronesian and Papuan Languages*. Canberra: Pacific Linguistics.



Project 6

Information structure in language acquisition

Project coordinator Christine Dimroth

Project members Bettina Braun, Aoju Chen, Christine Dimroth, Miriam Ellert, Juhani Järvikivi, Anke Jolink, Wolfgang Klein, Leah Roberts, Laura de Ruiter, Sarah Schimke, Goran Stanic, Giusy Turco, Josje Verhagen



Goals of the project

This project studies the linguistic realisation of information structure and its acquisition by children and adults in a variety of languages. The project examines the development of the relationship between semantic/pragmatic functions (topic, focus, given, new, contrast) and their corresponding formal devices. Among the various means which are typically used to encode information structure, priority will be given to the study of word order, intonation, pronominals, and particles. Project members use a variety of methods to explore these questions in learner data from different languages, including the investigation of corpora of spontaneous production data, elicited production, reaction time and eye-tracking techniques.

Subproject 1: Intonation encoding of topic and focus in child Dutch

In many languages, focus is typically realised with more intonational prominence than topic. In languages like Dutch, intonational prominence is primarily implemented via phonological means, i.e., placement of accent and choice of accent type (e.g., a weak rise

vs. a fall). In the case of identical accent types, acoustic prominence can differ in terms of gradient phonetic cues (e.g., duration and pitch excursion). In this project, Aoju Chen investigated how children acquire the intonational encoding of topic and focus in Dutch on the

basis of both spontaneous and elicited production.

With respect to phonological encoding, we found that children start out by marking focus mostly with a fall (annotated H*L) and topic with a downstepped fall (annotated !H*L) or complete devoicing in the late two-word stage. The use of devoicing and frequent use of !H*L instead of producing no accentuation in topic can be explained by poor control over pitch in very young children. At the age of 4 or 5, children lower pitch like adults when not accenting and become adult-like in accenting focus regardless of its position and sentence-initial topic but not accenting sentence-final topic. Their choice of accent type is however not adult-like in sentence-final focus in that they show no preference for H*L over !H*L. Such a preference for H*L is only developed at the age of 7 or 8. It is argued that the intonational variability in sentence-final focus in adult Dutch poses difficulties for the younger children in figuring out the preferred accent type.

With respect to the phonetic encoding, analysis of sentence-initial topic and focus accented with H*L shows that children use neither duration nor pitch-related cues to distinguish topic and focus at the age of 4 or 5, and use pitch-related cues but not duration for this purpose at the age of 7 or 8 (see Figure 1). To conclude, these results indicate that learning the intonation encoding of topic and focus is a gradual process and that phonological encoding is acquired earlier than phonetic encoding in Dutch.

The verb category of finiteness is typi-

Subproject 2: The acquisition of finiteness

cally realised by a particular inflection. But it is closely connected to numerous syntactic, semantic, and pragmatic properties of a sentence. Thus, its acquisition in a second language (L2) is a highly complex process. On the formal side, L2 learners have to learn, among other things, that verbs carry inflections for person and number and may move to a position reserved for finite verbs. On the functional side, they have to associate finiteness marking with the expression of assertion.

In two current PhD projects (Sarah Schimke, Josje Verhagen), the acquisition of finiteness in Turkish L2 learners of French, German and Dutch as well as Moroccan learners of Dutch was investigated at the formal and functional level. Participants were learners who acquired the target language in an immersion setting and were at various stages in the development of finiteness. A variety of methods was used to test both production and comprehension or processing of finite and non-finite utterances (i.e., elicited-production, elicited-imitation, picture selection, reaction-time experiments).

The data showed that the acquisition of finiteness is a gradual process, characterised by a shift from general semantic to morphosyntactic principles determining utterance organisation. Auxiliaries appeared to play a crucial role in this shift in German and Dutch, but not in French. Moreover, while the acquisition of formal finiteness marking could

be shown to coincide with the acquisition of the meaning of assertion in German, this was not the case in French, where a dissociation between formal and functional aspects of finiteness was observed: Learners acquire the formal properties of finiteness faster than in Germanic languages, but do not associate target-like meanings with these forms. In addition to target-language effects, clear influences of the source language were observed in the Turkish and Moroccan learners of Dutch. For all three target languages, however, it was found that learners show knowledge about finiteness in comprehension or processing tasks at a stage in which they are not yet able to use this knowledge in their production. For example, in a sentence-matching task, L2 learners of Dutch processed grammatically negated sentences with auxiliaries faster than ungrammatical ones, prior to the stage where they use auxiliaries in their production.

Selected publications

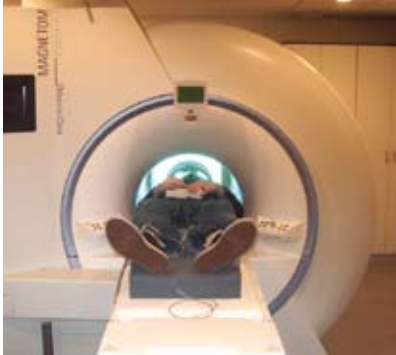
- Chen, A.** (2007). Intonational realisation of topic and focus by Dutch-acquiring 4- to 5-year-olds. In J. Trouvain & W.J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1553-1556). Dudweiler: Pirott.
- Klein, W.** (2008). The topic situation. In B. Ahrenholz, U. Bredel, W. Klein, M. Rost-Roth, & R. Skiba (eds.), *Empirische Forschung und Theoriebildung* (pp. 287-305). Berlin: Peter Lang.
- Narasimhan, B., & Dimroth, C.** (2008). Word order and information status in child language. *Cognition*, 107(1), 317-329.
- Roberts, L., Gullberg, M., & Indefrey, P.** (2008). On-line pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*, 30(3), 333-357.
- Verhagen, J., & Schimke, S.** (2009). Differences or fundamental differences? *Zeitschrift für Sprachwissenschaft*, 28, 97-106.

Other subprojects

- Pronoun resolution in first and second language learners.
- Intonation in reference tracking in second language learners.
- Word order in child language.
- Acquisition of givenness intonation in German.
- L2 learners' production and comprehension of referential continuity.
- The expression of additive and contrastive relations in L2 discourse.
- Scope particles and perspective taking in advanced second language discourse.
- The effects of fronting.
- The role of intonation and syntax in focus interpretation in children and adults.
- The interface between word order and intonation in focus marking in child language.
- The acquisition of focus in child Serbian.



Figure 1: Focus marking. Production elicited from L1 learners of Dutch



Project 7 Unification

Project coordinator Peter Hagoort

Project members Petra van Alphen, Giosuè Baggio, Jana Basnakova, Marcel Bastiaansen, Jos van Berkum, Ines Bramao, Dannie van den Brink, Dan Dediu, Christian Forkstam, Caroline Junge, Peter Hagoort, Miriam Kos, Laura Menenti, Karl-Magnus Petersson, Judith Pijnacker, Katrien Segaert, Irina Simanova, Tineke Snijders, Ana Carolina Sousa, Cathelijne Tesink, Julia Udden, Meinou de Vries, Lin Wang, Kirsten Weber, Roel Willems, Huadong Xiang



Goals of the project

The central question of this project is: **How are different sources of information which are retrieved from memory or provided by sensory input unified with language into an interpretation (comprehension) or message (production) beyond the single word level? Which neural networks are recruited for these unification operations? To what degree are these shared between production and comprehension, and what is the nature of their dynamic interplay with memory components, such as the mental lexicon or episodic memory of the prior discourse?**

The neural networks that support this unification process are under investigation. In a resting state study we found that the left language cortex shows a topographic organisation, where left inferior frontal cortex is connected to parietal and temporal areas in the left hemisphere according to information types (phonological, syntactic, semantic). Such a topographic pattern of connectivity is not observed in the right hemisphere. In addition, in fMRI studies we found that areas in frontal cortex have a modulatory role on the memory areas in temporal cortex. In this way we were able to characterise the dynamic interplay between different neural nodes in the brain's language network. Patients with autistic spectrum disorder were found to have an abnormal activation pattern, with a stronger reliance on the right frontal cortex to compensate for reduced functioning of the left frontal cortex.

Subproject 1: The neural infrastructure for the retrieval and unification of syntactic structure in sentence comprehension

In this subproject we investigate the division of labour in perisylvian language cortex between the left inferior frontal component and left temporal areas. This subproject brings together two different lines of research. One is related to models of parsing: the claims tested are derived from a computationally explicit, lexicalist model of parsing. This model assumes that with every incoming word one or more lexical frames are retrieved from the mental lexicon. These lexical frames are elementary syntactic trees, specifying the possible structural environment of the particular input word. Lexical frames that are retrieved

from memory will enter the unification space in a sequential fashion, as new words arrive as input. Then, in unification space binding operations between lexical frames are performed, resulting in an incremental structural interpretation of the sentence. During the unification process lexical frames are linked, and agreement features (number, gender, person, etc.) and word order constraints are applied. The other line of research relates to the neural infrastructure for language. Based on neurobiological arguments and meta-analyses of neuroimaging data, it is hypothesised that temporal cortex subserves memory for language,

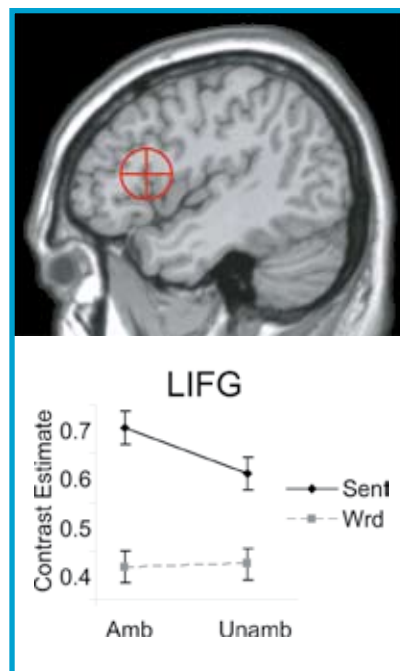


Figure 1 Red circle indicates the area in Left Inferior Frontal Gyrus (LIFG) in which an interaction (bottom) in the pattern of brain activation is observed between the type of language input (sentences vs. word strings) and the syntactic ambiguity (present or absent).

whereas unification operations require the contribution of left frontal cortex. This resulted in a strongly hypothesis-driven fMRI study. In this study participants read sentences and word sequences containing word-category (noun-verb) ambiguous words at critical positions. Regions contributing to the syntactic unification process should show enhanced activation for sentences compared to words, and only within sentences display a larger signal for ambiguous than unambiguous conditions (see Figure 1). The posterior LIFG (Left Inferior Frontal Gyrus) showed exactly this predicted pattern, confirming our

hypothesis that LIFG contributes to syntactic unification. The left posterior middle temporal gyrus was activated more for ambiguous than unambiguous conditions (main effect over both sentences and word sequences), as predicted for regions subserving the retrieval of lexical-syntactic information from memory. The dynamic interplay between these areas could be shown with the help of a psychophysiological interaction analysis.

Subproject 2: (Re)computing discourse models

While syntactic reanalysis has been extensively investigated in psycholinguistics, comparatively little is known about reanalysis in the semantic domain. We used event related brain potentials (ERPs) to keep track of semantic processes involved in understanding short narratives, such as, 'The girl was writing a letter when her friend spilled coffee on the paper'. We hypothesised that these sentences are interpreted in two steps: (1) when the progressive clause is processed, a discourse model is computed in which the goal state (a complete letter) is predicted to hold; (2) when the subordinate clause is processed, the initial representation is recomputed to the effect that, in the final discourse structure, the goal state is not satisfied. Critical sentences evoked larger Sustained Anterior Negativities (SANs) compared to controls, starting around 400 ms following the onset of the sentence-final word, and lasting for about 400 ms. The amplitude of the SAN was correlated with the frequency with which participants, in an offline probe-selection task, responded

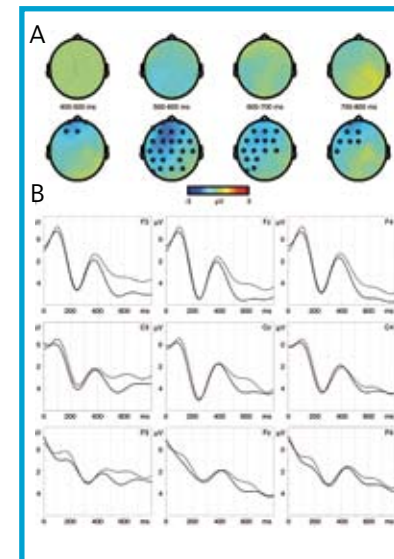


Figure 2 ERP waveforms for the sentence-final verbs. (A) Grand-average ($N = 24$) topographies displaying the mean amplitude difference between the ERPs evoked by the sentence-final verb in disabled compared to neutral accomplishments. Circles represent electrodes in a significant cluster. (B) Grand-average ($N = 24$) ERP waveforms from frontal, central and parietal electrode sites time locked to the onset (0 ms) of the verb in neutral and disabled accomplishments. Negative values are plotted upward.

that the goal state was not attained. Our results raise the possibility that the brain supports some form of non-monotonic recomputation to integrate information, which invalidates previously held assumptions (see Figure 2).

Other subprojects

- ERP studies on the interface between semantic and syntactic unification.
- Unification disorders in autism and Asperger.
- A brain computer interface for speaking.
- Artificial grammar learning.
- Syntactic priming and repetition suppression/enhancement.
- Oscillatory brain dynamics of language processing.
- Connectivity within the language cortex.
- Genetic modulations of the endophenotype for unification.
- Information structure and unification.

Selected publications

- Hagoort, P., & van Berkum, J.J.A. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society. Series B: Biological Sciences*, 362(1481), 801–811.
- Baggio, G., van Lambalgen, M., & Hagoort, P. (2008). Computing and recomputing discourse models: An ERP study. *Journal of Memory and Language*, 59(1), 36–53.
- Hagoort, P. (2008). The fractionation of spoken language understanding by measuring electrical and magnetic brain signals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 1055–1069.
- van Heuven, W.J.B., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex*, 18, 2706–2716.
- Koten, J.W.J., Wood, G., Hagoort, P., Goebel, R., Propping, P., Willmes, K., & Boomsma, D.I. (2009). Genetic contribution to variation in cognitive function. *Science*, 323(5922), 1737–1740.



Project 8

Phonological learning for speech perception

Project coordinator Anne Cutler

Project members Anne Cutler, Mirjam Ernestus, Martijn Goudbeek, Alexandra Jesse, Elizabeth Johnson, Caroline Junge, Valesca Kooijman, James McQueen, Holger Mitterer, Dennis Norris, Takashi Otake, Matthias Sjerps, Annelie Tuinman, Anita Wagner



Goals of the project

This project (which ran from 2001 to 2008) investigated how phonological structure is learned, deployed in understanding speech, and adapted to make speech perception more efficient. The three major research lines concerned learning a first phonology (in infancy); learning a second phonology (adult L2 learning); deploying and adapting the first phonology in speech perception.

Subproject: Perceptual retuning

The perceptual retuning of phoneme categories of the first phonology (Norris, McQueen & Cutler, 2003) was an early discovery of the project. In two-part experiments, listeners are first exposed to an ambiguous sound, in disambiguating lexical context. For instance, a sound halfway between /f/ and /s/ might occur in *horse* (so it should be interpreted as /s/) or in *giraffe* (so it would be /f/). Subsequent categorisation tests reveal that listeners adjust their categories to embrace the unusual sound. The adjustment requires very little exposure input, and generalises across the lexicon; it is an efficient mechanism for adapting to new talkers and speeding lexical recognition.

Matthias Sjerps' MSc research again confirmed the rapidity and generalisation of retuning, and discovered that a substituted sound can induce it as effectively as an ambiguous sound. Further explorations of retuning are summarised in Figure 1. Panel (a) shows the shift in Dutch listeners' categorisations of sounds varying from [ɛf] to [ɛs], following exposure to an ambiguous sound replacing either /f/ in words like *archieff* 'archive' or /s/ in words like *matras* 'mattress'. The midpoint [ʔ] represents the ambiguous sound actually used, but a

shift can be seen across the continuum: the /f/ category has expanded for listeners who heard [ʔ] as /f/, while /s/ has expanded (and /f/ correspondingly shrunk) for those who heard [ʔ] as /s/. These data are from investigations of whether phonemic retuning is position-specific; as in the 2003 study, the information driving the retuning here is lexical identity.

But lexical information is not essential; any reliable source of information will do. Panel (b) shows a similar shift in English listeners' categorisation functions for *flar/slar*, following exposure to an ambiguous sound in nonword-initial clusters in which [ʔ] was followed by /r/ or by /n/. In English, *fr-* is a possible onset (*fry, frog*) but *sr-* is impossible; conversely, *sn-* is an onset (*snip, snack*) but *fn-* is not. Even though the items are nonwords (*frul, snud* etc.), the sequence constraints force interpretation of [ʔ] as /f/ before /r/, but as /s/ before /n/; this interpretation then induces retuning.

Phonemes are not the only speech sounds which can be adjusted in this way, either. In languages with lexical tone, tone categories play a role in lexical access. Panel (c) shows Mandarin Chinese listeners' tone categorisations

after having been exposed to an ambiguous contour, between tone 1 and tone 2, in words requiring either tone 1 or tone 2; [ʔ] here represents the ambiguous tone used in exposure. It can be seen that again the whole tone category boundary has shifted, with the effect of expanding whichever tonal category [ʔ] was assigned to in exposure.

Finally, non-speech categories can also be retuned, both within the language domain (e.g., printed letters; Annual Report, 2002), and outside it. Panel (d) shows colour categorisation decisions about socks varying in hue from yellow

to orange. Here, '?' represents a hue between yellow and orange; in an initial exposure phase it was presented on, for instance, a carrot (causing [ʔ] to be categorised as orange) or on a banana (yellow). Again, participants shifted their entire category boundaries, expanding the colour category appropriate to '?' in the exposure phase. Category boundary flexibility, and the adaptation of category decisions to fit the constraints of the current perceptual situation, appears to be a powerful and general learning mechanism, allowing us to deal with complex signals which arrive rapidly, overlap, and vary with context. Speech

Selected publications

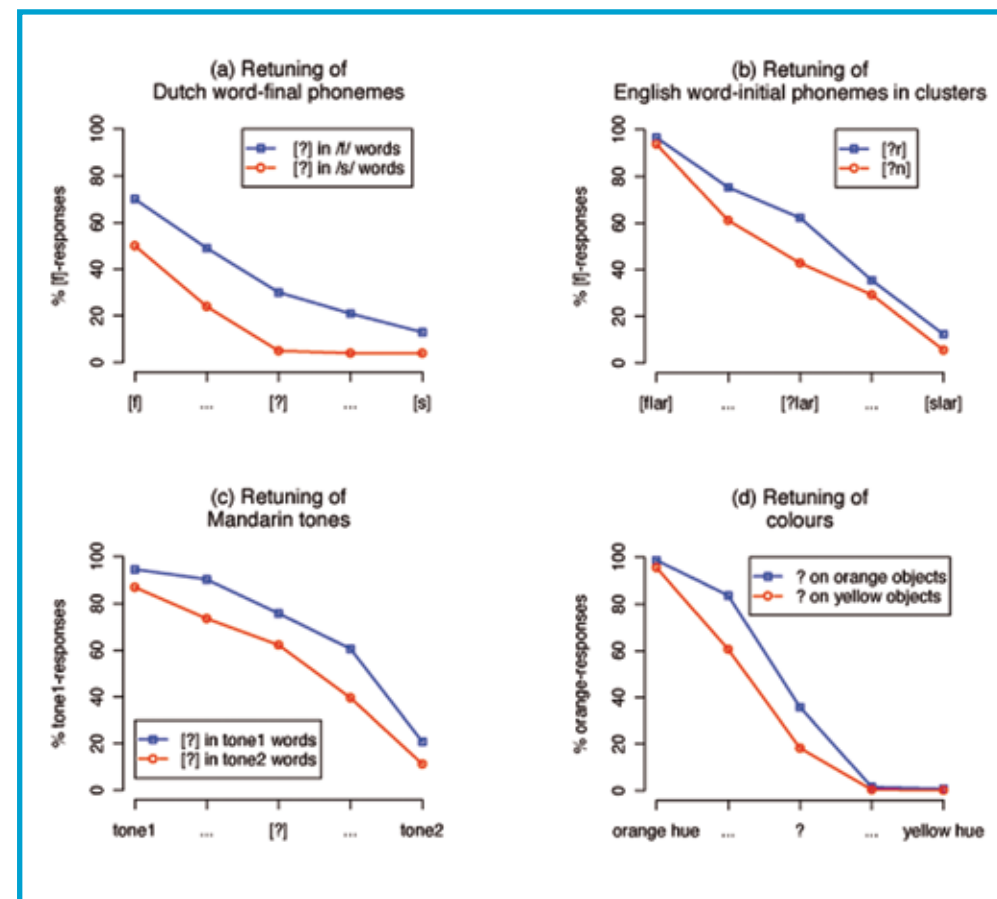
- Dietrich, C., Swingle, D., & Werker, J.F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences*, 104, 16027–16031.
- Snijders, T., Kooijman, V., Cutler, A., & Hagoort, P. (2007). Neurophysiological evidence of delayed segmentation in a foreign language. *Brain Research*, 1178, 106–113.
- Cutler, A. (2008). The abstract representations in speech processing (34th Bartlett Lecture). *Quarterly Journal of Experimental Psychology*, 61(11), 1601–1619.
- Mitterer, H., & de Ruiter, J.P. (2008). Recalibrating colour categories using world knowledge. *Psychological Science*, 19(7), 629–634.
- Murty, L., Otake, T., & Cutler, A. (2007). Perceptual tests of rhythmic similarity: I. Mora rhythm; Kim, J., Davis, C. & Cutler, A. (2008). Perceptual tests of rhythmic similarity: II. Syllable rhythm. *Language and Speech*, 50, 77–99, and 51, 342–358.

is one such signal, but it is not the only one. (Colour categories, of course, can be seen as semantic categories; see the report of project 5 for a further step in understanding how category retuning works.)

Other research

The PhD theses of Martijn Goudbeek (acquisition of auditory categories), Valesca Kooijman (electrophysiology of speech segmentation in infancy) and Anita Wagner (phoneme repertoire effects in speech sound processing) were defended in 2007–2008. In the infant speech perception laboratory, Kooijman found asymmetries between cortical and behavioural responses at the earliest stages of infant word recognition, and Caroline Junge observed a correlation of later vocabulary development with the presence of a cortical

reflection of speech segmentation at 10 months. Caretakers align co-speech gestures to the prosody in speech to children, and audiovisual alignment suffices for two-year-olds to deduce word meaning. Artificial-language learning studies with English, Dutch, and French listeners revealed both cross-language similarities and differences in the use of different acoustic cues to segmentation of continuous speech into words. Second-language listening studies included further investigations of the source of the difficulty of L2 speech perception in noisy environments, a demonstration of greater sensitivity to prosodic goodness by non-native than by native listeners, and the finding by Annelie Tuinman that casual-speech processes in L2 can reduce lexical processing efficiency, if these processes are unfamiliar because they do not occur in the L1.





Project 9

Language documentation and data mining



Goals of the project

As part of the Institute's wider work on cross-linguistic patterns of language coding and language use (see projects 4 and 5), primary research is carried out on many languages for which little prior information exists. These languages are often spoken by only a few hundred or thousand people in undeveloped parts of the world. There is an increasing urgency to document these small languages, which are disappearing at the rate of about one per week, and which represent most of the world's linguistic diversity (approximately 80% of languages are spoken by less than 100,000 people). With their disappearance, we not only lose millennia of cultural evolution, but also important scientific evidence about the range of human language variation.

acterised by complex verbal derivation, compounding and incorporation. The traditional hunter-gatherer mode of subsistence is reflected in various properties of language use, including an unusual pattern of question-answer usage (see p. 16), and a reliance on an absolute spatial system, which has been the focus of sustained psychological work by Christian Rapold and Daniel Haun.

ple, is being done on Lanoh. The collection of word lists has allowed the inference of a detailed family tree linking the languages, using bioinformatic methods (see p. 32). Detailed work on one of these languages, Jahai, has raised a number of interesting issues for other institute projects, by showing for example that covert macro-structures ('semplates') in the lexicon can impose structure on a number of apparently unrelated lexical sets.

Data from languages like these often introduce new, unexpected features which prove quite challenging for current theories about language universals. But we can also be sure that the languages have interesting properties that we have not yet detected. For that reason, it is important to archive as rich a record of these languages as possible for posterity. Language documentation yields mountains of data – the institute currently has about 48 Terabytes of endangered language data. The institute's technical group, with funding from the Volkswagen Foundation, the EU and NWO, has built the technical infrastructure Language Archiving Technology (LAT) to store and navigate through these records – to annotate video records, link and display online multimedia dictionaries, and retrieve precise locations in the data from metadata or annotations (see p. 34). The integration of data from languages all around the world, including those from the DoBeS programme from the Volkswagen Foundation makes possible entirely new methods of research into the diversity of languages and the nature of language universals. The institute has done much primary research in the area



Figure 1 Field sites where 'small' languages are under investigation.

Currently, work proceeds in twenty field sites around the world. The research involves language description from the ground up: developing a writing system, producing grammars and dictionaries, and the collection of video- and audio-tape so that it can be transcribed and annotated as a permanent record in the institute's archives. In addition, special elicitation and experimentation is carried out for this and other projects, including doctoral thesis research.

Subproject 1

We give two examples of the larger scale projects (see p. 27 for complete list). One of these projects (Thomas Widlock, Gertie Hoymann, Christian Rapold) is documenting †Akhoe Haillom, a so-called Khoisan language in Namibia spoken by just 600 hunter-gatherers. It has a complex sound system, with clicks and other unusual consonants, and tones, while the grammar is char-

Subproject 2

Another project (Niclas Burenhult, Claudia Wegener, Neele Becker) is focussed on the Semang, hunter-gatherers of Malaysia (closely related to the Semai, researched by Sylvia Tufvesson). Survey work is carried out to establish the historical relations between the seven main linguistic groups in Malaysia and Thailand, who speak Aslian languages. The first detailed descriptive work, for exam-

Documentation projects

- 1 Yuraké (*Bolivian isolate*): Rik van Gijn, Sonja Gipper (*PhD student*); Vincent Hirtzel (*College de France*): Funded by the Volkswagen Foundation.
- 2 †Akhoe Haillom (*Khoisan, Namibia*): Thomas Widlok (*Radboud U. @ MPI*), Gertie Hoymann (*PhD student*), Christian Rapold (*Zürich*). Funded by the Volkswagen Foundation.
- 3 Semang (*Six Aslian languages*): Niclas Burenhult, Claudia Wegener, and Neele Becker. Funded by the Volkswagen Foundation.
- 4 Five languages of Cape York (*Australia*): Clair Hill (*PhD student*), Jean-Christophe Verstraete (*U. Leuven*), Peter Sutton (*Adelaide*), Alice Gaby (*Berkeley*). Funded by Hans Rausing Endangered Languages Project.
- 5 Yéli Dnye (*Papua New Guinea*): Stephen Levinson, Penelope Brown.
- 6 Kariil (*Austroasiatic, Laos*) & Lao: Nick Enfield.
- 7 Savosavo (*Solomons*): Claudia Wegener (*PhD student*).
- 8 Rotokas (*Bougainville, PNG*): Stuart Robinson (*PhD student*).
- 9 Tzeltal (*Mayan, Mexico*): Penelope Brown.
- 10 Siwu (*Ghana*): Mark Dingemans (*PhD student*).
- 11 Kata Kolok (*Balinese sign language*): Connie de Vos (*PhD student*).
- 12 Kilivila (*Papua New Guinea*): Gunter Senft.
- 13 Zapotec (*Mexico*): Mark Sicoli.
- 14 Semai (*Malaysia*): Sylvia Tufvesson (*PhD student*).
- 15 Warlpiri (*Pama-Nyungan, Australia*): Carmel O'Shannessey (*PhD student*).
- 16 Lowland Chontal (*isolate, Mexico*): Loretta O'Connor (*U. Hamburg*) Funded by the Volkswagen Foundation.

Data mining teams

- 17 Sahul Project (*typology of New Guinea and Australia*): Michael Dunn, Ger Reesink, Ruth Singer, Miriam van Staden, Pieter Muysken and Stephen Levinson. Funded by NWO.
- 18 Marquesan Lexicon project: Gaby Cablitz, Jacquelyn Ringersma. Funded by the Volkswagen Foundation.
- 19 DoBeS exploitation tools: Peter Wittenburg, Claus Zinn, Paul Trilsbeek, et al. Funded by EU and the Volkswagen Foundation.

of New Guinea and Australia (including doctoral and postdoctoral work), complementing earlier research. This large data set has allowed the application of bio-informatic methods for extracting likely prehistoric relations between lan-

guages – from, for the first time, purely structural data. This type of work will be picked up as the major theme for the new research group, 'Evolutionary processes in language and culture' (see p. 32).

Selected publications

- Haun, D.B.M (2007). *Cognitive cladistics and the relativity of spatial cognition*. Doctoral thesis, Radboud University (http://webdoc.uhn.ru.nl/mono/h/haun_d/cognclant.pdf).
- Dunn, M., Levinson, S.C., Lindström, E., Reesink, G., & Terrill, A. (2008). Structural Phylogeny in Historical Linguistics: Methodological Explorations Applied in Island Melanesia. *Language*, 84(4), 710–759.
- Wegener, C. (2008). *A grammar of Savosavo: A Papuan language of the Solomon Islands*. Doctoral thesis, Radboud University Nijmegen.
- Software tools downloadable from: <http://www.mpi.nl/tools/>

Max Planck Research Group Communication before language



Goals of the project

Human communication is premised on an understanding of others' minds and cooperative motives for acting together. How do these abilities emerge, and how do infants communicate before they have language? We investigate infants' developing social cognition and social motivation in relation to their emerging prelinguistic communication within their social and cultural contexts. Our work is motivated by the two central ideas that (a) the psychological basis of human communication develops ontogenetically prior to language and can be first expressed in gestures; and (b) social and cultural differences in interaction may influence infants' emerging prelinguistic communication skills.

Subproject 1: Infant social cognition and joint action

To investigate the psychological foundations of human communication, we conduct lab-based experimental studies of infants' understanding of other persons' actions and epistemic states, their comprehension of referential intentions, and their understanding and initiation of joint actions. In a new interaction paradigm we examine whether prelinguistic infants anticipate a person's erroneous action and use their social and cognitive communication skills to correct false belief. Using object-choice tasks, we test the conditions under which infants understand gestural reference, in particular to occluded and displaced referents. In a new imitation paradigm, we investigate whether infants encode the cooperative nature of others' joint actions, such that they later attempt to reproduce these observed actions in a cooperative manner. In eye-tracking studies, we test whether infants attribute to agents preferences for individuals who have previously acted jointly with them.

Subproject 2: Socialisation of prelinguistic communication

In this subproject we conduct longitudinal and cross-cultural observation studies on infants' interactions with their social environment. For example, we compare the frequency of infants' naturally occurring social interactions in a Yucatec Mayan village in Mexico with that of infants in the Netherlands. We look for similarities and differences in infants' interaction styles and in the input they receive. Further, we investigate whether the pointing gesture – an index of infants' prelinguistic referential communication – emerges around the same ages across different cultures (e.g., Bali, Japan, Mexico, Rossel Island in Papua New Guinea, Peru, India). In a longitudinal study in the Netherlands, we investigate individual and developmental differences in the emergence of prelinguistic gestures with regard to caregivers' interactions. In intervention studies we aim at testing when and to what extent training has an effect on

Group coordinator Ulf Liszkowski
Group members Christine Fawcett, Birgit Knudsen, Daniel Puccini, Dorothé Salomo

Selected Publications

Liszkowski, U. (2008). Before L1: A differentiated perspective on infant gestures. *Gesture*, 8(2), 180–196.
Liszkowski, U., Schäfer, M., Carpenter, M., & Tomasello, M. (in press). Prelinguistic infants, but not chimpanzees, communicate about absent entities. *Psychological Science*.

infants' naturally occurring gestures and communication skills.

Other subprojects

- Colour categories in preverbal infants (in collaboration with project 5).
- Third-person watching dyadic conversations.





Max Planck Research Group Adaptive listening

Max Planck Research Group Comparative cognitive anthropology



Group coordinator

Andrea Weber

Group members

Adriana Hanulikova, Frank Eisner
(from 2009)

Goals of the project

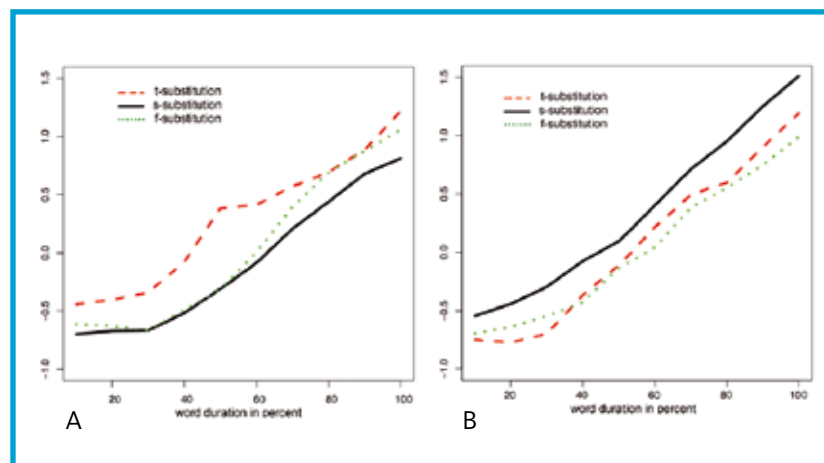
In June 2008, a new independent research group on Adaptive listening began its work. The goal of the group is a better understanding of the mechanisms and representations underlying the processing of foreign-accented speech. Foreign-accented speech reflects the sound structure of two languages, and listeners have varying experience with accents. Current research draws on these aspects to investigate cross-linguistically how the speech processing system adapts to accent-specific variation in the speech signal.

Subproject 1: Segmental deviations and lexical access

Mispronunciations are a conspicuous marker of foreign accents. This project explores how and when mispronunciations influence lexical activation; in particular, research focuses on the role of frequency of occurrence and accent-specificity of segmental deviations. In an English eye-tracking study, Hanulikova and Weber are investigating whether lexical activation is sensitive to accent-specific tendencies for mispronunciations. Although Dutch and German learners of English substitute th interchangeably with s, t, and f, Ger-

man speakers incline to s-substitutions, as in 'seft' for the English word theft, but Dutch speakers to t-substitutions; f-substitutions are perceptually most similar to th but occur only rarely. Eye movements of Dutch and German listeners show that s-, t-, and f-substitutions indeed activate corresponding th-words, but the degree of activation differs between groups, with activation being strongest for the dominant substitution of the listener's own accent (Figure 1). Together, the results suggest that mispronunciation frequency influences

lexical access, co-existing with the well-attested influence of perceptual similarity. Currently, we are testing whether brain responses show a similar sensitivity to the perceived difference between standard forms and mispronunciations. Furthermore, a number of cross-modal priming studies explore whether mispronunciations must be specific to the listeners' accent to cause spurious lexical activation. This includes investigating the role of fine phonetic detail and the transfer of activation patterns to related accents.



Other subprojects

- Perceptual adaptation in children at different development stages.
- Brain responses to syntactic violations in foreign-accented speech.

Figure 1 Mean fixations over time to a displayed th-word for trials with s-, t-, and f-substitutions, log odd scale: (A) Dutch listeners; (B) German listeners

Goals of the project

The research group for Comparative cognitive anthropology (launched in August 2008) is a joint effort between the Max Planck Institutes for Psycholinguistics and Evolutionary Anthropology. It aims to understand both constancies and variation in human cognition and their sources in primate inheritance on the one hand, and cultural diversity on the other. It also aims to investigate the psychological mechanisms behind uniquely human cultural accumulation and 'speciation', such as individual and social learning, cooperation, conformity and norms.

Subproject 1: Recognising relational similarity

Recognition of relational similarity is the ability to understand that defining object properties might not lie in objects individually, but in the relations of the properties of various objects to each other, and also that objects sharing relations with their respective surroundings are similar to each other. This powerful individual learning mechanism is the basis of many skills such as understanding propositional structure and predication, inductive inference, categorisation and understanding analogy and metaphor. We have found the ability to recognise relational similarities in five species of

great apes, including human children. We have discovered that children above 4 years of age, bonobos and chimpanzees, but not younger children, gorillas, and orangutans, display some ability to reason by non-causal relational similarity. We conclude that recognising relational similarity is not in its entirety unique to the human species.

Subproject 2: Recognising being imitated

Human infants imitate not only to acquire skills, but also as a fundamental part of social interaction. Reciprocal imitation creates social affiliations between humans, which play a crucial role in quick, automatic in-group formation amongst strangers as well as automatic maintenance of affiliations between friends. Infants demonstrate recognition of being imitated by engaging in so-called testing behaviours. This behaviour affords the ability to recognise structural and temporal contingencies between actions across agents and the ability to understand the directional im-

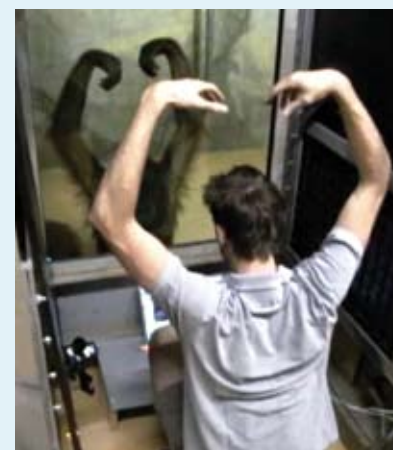


Figure 1 Experimenter and female orangutan (*Pongo pygmaeus*) interacting in an experiment investigating imitation recognition in non-human great apes.

Group coordinator

Daniel Haun

Group members

Alenka Hribar, Yvonne Rekers, Nadja Richter, Disa Sauter, Sebastian Schütte

Selected Publications

- Haun, D.B.M., & Call, J.** (2008). Imitation recognition in great apes. *Current Biology*, 18(7), R288–R290.
- Haun, D.B.M., & Call, J.** (2009). Great Apes' capacities to recognise relational similarity. *Cognition*, 110(2), 147–159.

pect of one's own actions on others' actions. In the past year, we investigated imitation recognition in four species of non-human great apes. Our results show a general prevalence of imitation recognition in all great apes. Our next step is to investigate whether imitation prompts human-like social consequences in other great ape species.

Other subprojects

- Time.
- Space.
- Emotional expression across cultures and species.
- Conformity in humans and other apes.
- Norms of sharing and fairness across cultures.



Max Planck Research Group Evolutionary processes in language and culture

Max Planck Fellow Syntax, typology and information structure



Group coordinator Michael Dunn

Goals of the project

This independent research group investigates language diversity and change as part of an integrated evolutionary system. It takes an evolutionary perspective, borrowing computational tools from genetics and biology, and integrates probabilistic, quantified approaches to phylogenetics with rigorous tests of different models of the interaction between elements of language, contact and geography, and cultural variation. This group was awarded in 2008, and is currently in the startup phase.

Subproject 1: Correlated evolution of language structure

What constrains the variation of language? We use modern computational methods to answer this question from a broad phylogenetic perspective. Linguistic typology is currently struggling with the problem of making statistically valid statements about non-independent observations of typological variation: Are the structures in two languages similar because this is how language is, or is it simply because these two languages descend from a common ancestor? Bio-informatic methods can be used to model evolutionary processes in language to make statistically sensitive, probabilistic measurements of typological variation within a phylogenetic tree. This will let us make statistical statements about typological stability, and to measure subtle typological interdependencies.

to infer historical relationships between languages. Bayesian methods require a great deal of computational power to carry out, but very much enrich the historical insight that can be gained from linguistic data. They allow incorporation of realistic assumptions about the processes of language change, including geography, demography, cultural variation, and contact. Importantly, Bayesian phylogenetic inference is tolerant of the 'noisy' data, and allows quantification of uncertainty and the investigation of conflicting phylogenetic signals.

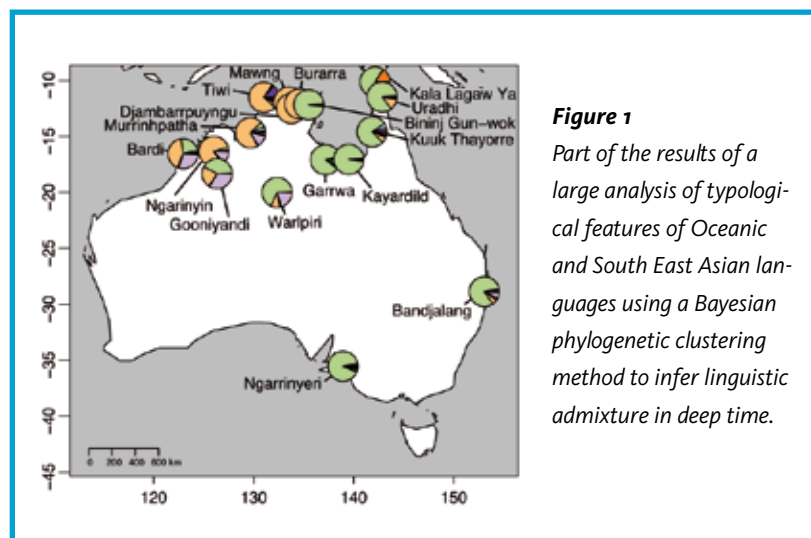
Selected Publications

Dunn, M., Foley, R., Levinson, S.C., Reesink, G., & Terrill, A. (2007).

Statistical reasoning in the evolution of typological diversity in Island Melanesia. *Oceanic Linguistics*, 46(2), 388-403.

Hunley, K.L., Dunn, M., Lindström, E., Reesink, G., Terrill, A., Healy, M., Koki, G., Friedlaender, F., & Friedlaender, J.F. (2008).

Genetic and Linguistic Coevolution in Northern Island Melanesia. *PLoS Genetics*, 4(10), e1000239.



Subproject 2: Language phylogenies

Understanding the history of language families as a whole is a prerequisite for modelling the evolution of features of individual languages. We supplement traditional historical linguistics with newer Bayesian phylogenetic methods

Goal of the project

This project began in October 2008. Its focus is the interaction between discourse pragmatics and syntax in languages of a range of structural types. This interaction happens on several levels and can potentially affect grammar in various ways. On the phonological level one of the virtually universal signals of the discourse pragmatic status of a constituent is its prosodic prominence or lack thereof: one way that speakers mark information that is assumed to be shared with interlocutors is prosodically de-emphasising it, whereas they typically mark the important information being introduced in the sentence by prosodically emphasising it. On the morphological level, for example, some languages have particles which signal that a constituent is the topic expression or that it is the focus of an assertion or question, while case assignment in some languages may be directly sensitive to discourse status rather than just signalling syntactic or semantic relations.

On the syntactic level, there are manifold influences, ranging from constraining word order to motivating the displacement of syntactic constituents, from creating units central for certain constructions to constraining core syntactic operations like subject selection. These interactions differ from language to language, e.g., in some discourse pragmatics strongly affects word order in sentences (e.g., Russian), while in others it is a much smaller factor (e.g., English), in some languages subject selection can be strongly influenced by the discourse pragmatic status of an argument (e.g., Tagalog, Swedish), whereas in other it plays no role at all (e.g., Lakota, Warlpiri). Given this cross-linguistic variation in the manifestation of information structure, an important question arises: what are the co-occurrence patterns of these interactions? That is, in a given grammatical system, only some of these interactions occur. Are there constraints on the co-occurrence of these interactions? This leads to a larger question: what role does information structure play in explaining cross-linguistic differences in grammatical systems? The goal of this project is to explore these questions, which will be developed as subprojects. It is expected that there will be substantial interaction with other projects, especially with project 6.

Group coordinator Robert van Valin





Technical group

Group head Peter Wittenburg

Group members Eric Auer, Herbert Baumann, Peter Berck, Dik van den Born, Daan Broeder, Hennie Brugman, Ger Desserjer, Reiner Dirksmeyer, Alex Dukers, Matthias Egger, Micha Hulsbosch, Gerd Klaas, Alexander Klassman, John Nagengast, Jacquelin Ringersma, Albert Russel, Han Sloetjes, Marc Snijders, Paul Trilsbeek, Dieter van Uytvanck, Tobias van Valkenhoef, Kees van der Veer, Ad Verbunt, Rick van Viersen, Johan Weustink, Menzo Windhouwer, Florian Wittenburg, Peter Wittenburg, Niclas Wood, Claus Zinn

Goals of the group

The Technical group (TG) has two major goals: to provide the infrastructure of labs, servers and field equipment for the day to day running of the institute, and to develop new systems and software that enable new scientific developments both within the institute and for the broader scientific community.

Computer systems

The institute's server systems have been substantially upgraded to provide state of the art computing and storage. These systems have to be able to handle large data flows from labs, enable fast access to central Max Planck supercomputers, and send data to remote sites for backup. The institute has played a role in shaping a new grid for accessing eResources within the Max Planck Society.

Experimental labs

The institute has built and maintains 6 reaction time labs, 4 eye movement labs, 1 Faraday-caged ERP lab, 1 gesture lab as well as 1 babylab on campus. In addition, a virtual reality lab is under construction.

The main neuroimaging facility, including infant EEG and eye-tracking labs, is housed in the Donders Institute for Brain, Cognition and Behaviour where 1.5t and 3t fMRI, MEG and EEG labs are maintained by a separate technical group in close coordination. During the last year the experimental systems have been harmonised between the two facilities.

Field expeditions

The institute makes use of over 20 field-sites around the world which allow cross-linguistic comparison and experiment.

During the period of review, 42 field trips were fully equipped with everything from solar panels to portable eye-trackers. Much specialised equipment has to be acquired, tested and adapted. This experience is passed on to the endangered language community, which now makes extensive use of the advice and tools for collecting and archiving field data.

Archive management

The digital archive now contains about 48 Terabytes of data in a well-organised, consistent, and accessible manner with a current annual increase of 12 TB. This material includes a great deal of endangered language recordings, constituting an invaluable resource for future data-driven research, making the institute

the major repository in the world for this kind of data (see p. 26). The holdings include the VW-funded endangered language corpora (DoBeS), the institute's own substantial 'exotic' language holdings, major language acquisition corpora (from Russian to Tzeltal), and material from other Max Planck Institutes or sources, such as ethnographic material collected by the Eibl-Eibesfeldt group. These endangered language corpora are served by a world-wide network of servers with the institute's archive in the centre (see Figure 1), which make it practical both to deliver the resources back to the countries of origin and to back up the indigenous language material at seven distinct sites. In addition to the digitisation of new and old recordings, much curation work was carried

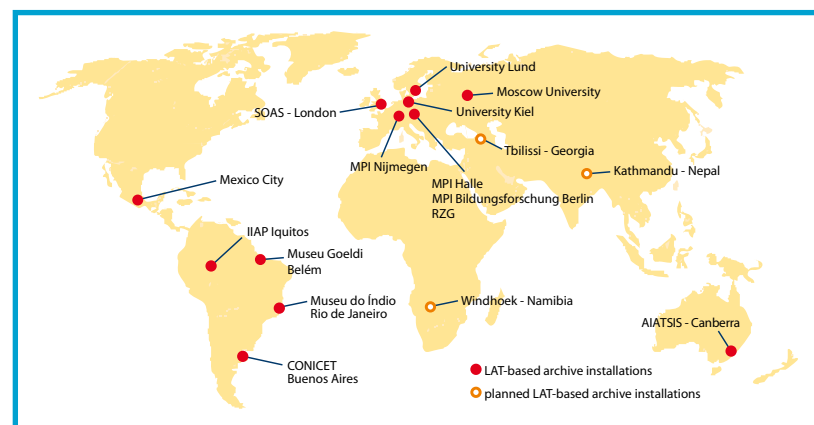


Figure 1 LAT-based archives for language resources.

out, such as transforming lexica into well-structured formats compliant with the ISO Lexical Markup Format.

The maintenance of the archive with good metadata for retrieval is quite demanding. In addition new systems for uploading and using the data need constantly to be devised. In 2006-8 an entire software suite has been created and is continuously being improved to 'ingest' new data, annotate and enhance existing records, and to access and search the stored data. New tools include web-based resources, allowing remote access, offline working, and the ability to link and explore data in creative ways. The institute is constantly improving its data mining tools, and has created a tool for representing indigenous knowledge of specific domains or 'conceptual spaces', with semantic links connecting video, stills, text, etc.

Language processing resources

As a general service to the language sciences, the group plays a key role in the multinational CLARIN project which works to establish a unified infrastructure and access mode to language resources, which are currently fragmented and in diverse technologies. The aim is to create an integrated 'observatory' of all language resources and tools and to enable a user-friendly, single sign-on access to these valuable resources. CLARIN builds on earlier European initiatives, in which the technical group also played a leading role.

Selected publications

- Ringersma, J., & Kemps-Snijders, M.** (2007). Creating multimedia dictionaries of endangered languages using LEXUS. *Proceedings of INTERSPEECH 2007* (pp. 1529-1532) [CD-ROM].
- Broeder, D., Declerck, T., Hinrichs, E., Piperidis, S., Romary, L., Calzolari, N., & Wittenburg, P.** (2008). Foundation of a component-based flexible registry for language resources and technology. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08)* (pp. 807-811) [CD-ROM].
- Kemps-Snijders, M., Klassmann, A., Zinn, C., Berck, P., Russel, A., & Wittenburg, P.** (2008). Exploring and Enriching a Language Resource Archive via the Web. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08)* (pp. 1433-1436) [CD-ROM].
- Trilsbeek, P., Broeder, D., Valkenhoef, T., & Wittenburg P.** (2008). A Grid of regional language archives. Exploring and enriching a language resource archive via the web. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08)* (pp. 1474-1477) [CD-ROM].
- Zinn, C., Ringersma, J., Cablitz, G., Kemps-Snijders, M., & Wittenburg, P.** (2008). *Constructing knowledge spaces from linguistic resources*. Proceedings of the CIL 18 Workshop on Linguistic Studies of Ontology: From Lexical Semantics to Formal Ontologies and Back, Seoul.



Events and activities



Workshops and conferences 2007				Workshops and conferences 2008			
January 26 February 28 - March 3	Australianist narrative workshop Workshop on functional elements: variation in learner systems at the 29 th annual conference of the German Linguistics Society (DGfS)	MPI Nijmegen Siegen, Germany	Hill Dimroth, and Jordens	February 14 - 16 March 11 March 17 - 19	Workshop on knowledge and responsibility in social interaction MesoSpace workshop CLARIN kick-off meeting	Århus U., Denmark MPI Nijmegen Nijmegen	Stivers Levinson Wittenburg, Van Uytvanck
March 12 - 16	Workshop on brain mechanisms and cognitive processes in the comprehension of discourse	Lorentz Center Leiden, Netherlands	Van Berkum, Ferstl, Van den Broek	March 27 - 28 April 11	eScience seminar on unique and persistent identifiers Workshop on expressives	MPI Nijmegen MPI Nijmegen	Wittenburg, Snijders Tufvesson, Enfield
March 16 - 23 March 22 - 23	Wenner-Gren symposium on ritual communication eScience seminar on service and information registries	Sintra, Portugal Göttingen, Germany	Basso, Senft Wittenburg, Snijders	April 14 - 18 May 07	Workshop on question-response sequences 12 th Berlin colloquium of the Gottlieb Daimler and Karl Benz Foundation 'Is beauty measurable?'	MPI Nijmegen MPI Nijmegen Akademie der Konrad-Adenauer- Stiftung, Berlin, Germany	Enfield, Stivers Klein
March 29 - 31	Workshop on information structure in adult and child language	MPI Nijmegen	Narasimhan, Chen Dimroth, Klein	May 14 June 01	Workshop on measuring vowel quality 3 rd workshop on the representation and processing of Sign Languages: construction and exploitation of Sign Language corpora	MPI Nijmegen Marrakech, Morocco	Chen Zwitzerlood
April 16 - 17 April 19	Workshop on Mon-Khmer linguistics: new developments Workshop on Birdsong, speech and language; converging mechanisms	MPI Nijmegen Utrecht, Netherlands	Burenhult, Enfield Hagoort, Bolhuis, Everaert	June 06 June 10 June 12 - 13 June 15 - 17	Workshop on human locomotion across languages Workshop on geocentric space DoBeS workshop 1 st Nijmegen workshop on speech reduction	MPI Nijmegen MPI Nijmegen MPI Nijmegen MPI Nijmegen	Majid, Kopecka Dasen, Senft Trilsbeek Ernestus, Warner (U. of Arizona)
April 21 - 24	Symposium on gesture and the acquisition of a second language, at the American Association of Applied Linguistics (AAAL) annual conference 2007	Costa Mesa (CA), USA	Gullberg, McCafferty	June 19	CLARIN PID workshop	MPI Nijmegen	Wittenburg, Van Uytvanck
May 10 - 11	Workshop on visual prosody in language communication	MPI Nijmegen	Jesse, Johnson, Zondervan Kempen	June 19 - 20 July 28	eScience seminar on aspects of long term archiving Symposium on the prosodic structure of Dutch one-word utterances: 'intonation or word-stress?' at the 6 th international congress for the study of child language (IASCL)	Göttingen, Germany Edinburgh, UK	Wittenburg, Snijders Fikkert, Chen
June 08	Session on neurocognition of unification in natural-language processing, during the 6 th Dutch endo-neuro-psycho meeting	Doorwerth, Netherlands	Trilsbeek	September 12	Thematic panel at the 18 th annual conference of the European Second Language Association (EUROSLA)	U. of Aix-en- Provence, France	Dimroth, Benazzo
June 14 - 15 June 20 - 21 July 26 - 28	DoBeS workshop eScience Seminar on secure server and service systems 2 nd interdisciplinary Max Planck PhD net workshop on communication	MPI Nijmegen Munich, Germany Frankfurt a.M., Germany	Wittenburg, Snijders Herbst	September 22 - 26	Interspeech 2008 (9 th in the annual Interspeech series; 12 th in the series of SST conferences)	Brisbane, Australia	Cutler (member of organising committee)
August 27 - 31	Interspeech 2007 jointly organized by the speech science and technology community of the Netherlands and Belgium, under the sponsorship of the International Speech Communication Association (ISCA)	Antwerp, Belgium	Cutler (member of the organising committee)	October 03	Workshop on language change in bilingual communities. This workshop was part of the 23 rd Scandinavian conference of linguistics	Uppsala U., Sweden	Zwitzerlood
October 6 - 7	2 nd annual meeting of the German science foundation's priority programme SPP1234 'Phonological and phonetic competence: between grammar, signal processing, and neural activity'	MPI Nijmegen	Jesse, Mitterer	October 3 - 4	PhD student conference: Donders discussions 2008	Donders Institute Nijmegen	Menenti, Aarts, Helmich, Van Dongen FitzPatrick, Van Leeuwen Meulenbroek, Volman
October 08 October 11 - 13	4 th annual meeting of the phonetics and phonology (P&P4) Workshop on competition meets cooperation	MPI Nijmegen MPI and Radboud U., Nijmegen	Mitterer, Jesse Enfield, Stivers	October 6 - 9	CLARIN WP2/5 workshop	Berlin, Germany	Wittenburg, Van Uytvanck
October 12 - 13	Workshop on oscillatory neuronal dynamics in psycholinguistic research	MPI and Donders Institute Nijmegen	Bastiaansen, Hagoort	October 14 - 15 October 17	eScience seminar on metadata infrastructures Workshop on the language-specific outcomes of the question-response sub-project	Berlin, Germany MPI Nijmegen	Wittenburg, Snijders Enfield, Stivers
October 17 - 18	2 nd A. Guiora roundtable conference in the cognitive neuroscience of language: the cognitive and neural prerequisites for time in language	MPI Nijmegen	Gullberg, Indefrey, Klein, Hagoort, Ellis, Schumann	October 30 - Nov. 1 October 26 - Nov. 1	2 nd workshop on knowledge and responsibility in social interaction Landscape in language: a transdisciplinary workshop	Lyon, France Albuquerque (NM) Chinle (AZ), USA	Stivers, Mondada Burenhult, Levinson Mark & colleagues (SUNY, Buffalo)
October 25 - 26 October 27	eScience seminar on open achievable formats European Australianists workshop	Berlin, Germany MPI Nijmegen	Wittenburg, Snijders Singer	November 10 - 11	Workshop on web services architecture in CLARIN	Munich, Germany	Wittenburg, Van Uytvanck
November 23 - 24	Workshop on language processing in first and second language learners	MPI Nijmegen	Indefrey, Gullberg, Cutler	December 15 - 18	Workshop on how social actions are recognized and performed in social interaction	MPI Nijmegen	Enfield, Stivers
December 12 December 14 - 15	Workshop on subordination in South American languages 11 th Winter conference of the NVP (Dutch Psychonomic Association)	MPI Nijmegen Egmond aan Zee, Netherlands	Van Rijn Mitterer				

Events and activities

Nijmegen Lectures 2007

Constructions and the nature of generalisation in language

The 2007 Nijmegen Lectures were given from 10-12 December by Adele Goldberg (*Princeton U.*). The series included three morning lectures: 'The constructionist approach to language', 'Learning constructions', and 'Accounting for generalisations across languages'. Discussants in the afternoon seminars were Arie Verhagen (*U. Leiden*), Jan Wouter Zwart (*U. Groningen*), Michiel van Lambalgen (*U. of Amsterdam*), Rens Bod (*U. of Amsterdam*), Heike Behrens (*U. Basel*), Christine Dimroth (*MPI for Psycholinguistics*), Ewa Dabrowska (*U. of Sheffield*), Bettelou Los (*Radboud U. Nijmegen*) and Leon Stassen (*Radboud U. Nijmegen*). The lectures were organised in collaboration with the Radboud University Nijmegen by Mirjam Ernestus, Ad Foolen, Bhuvana Narasimhan and Nanjo Bogdanowicz.

Nijmegen Lectures 2008

The nature and origins of language: A genetic perspective

The Nijmegen Lectures 2008 were given from 10-12 December by Simon Fisher (*U. of Oxford*) and Gary Marcus (*New York U.*). The series included two morning lectures on each of the three days: 'The genetics of language?' (GM) and 'The language of genetics' (SF) on day 1, 'Bridging genes, brains, and language' (SF) and 'What is the language faculty made of? Evidence from human infants and molecular biology' (GM) on day 2, and 'Language evolution and the genomics revolution' (SF) and 'Language as kluge' (GM) on day 3. Discussants in the afternoon sessions were Katharina Dworzynski (*King's College London*), Andreas Papassotiropoulos (*U. of Basel*), Franck Ramus (*CNRS, Paris*), Kurt Hammerschmidt (*German Primate Center, Göttingen*), James Hurford (*U. of Edinburgh*), Constance Scharff (*Freie U. Berlin*), Bart de Boer (*Vrije U. Amsterdam*), Dan Dediu (*MPI for Psycholinguistics*), and Wolfgang Enard (*MPI for Evolutionary Anthropology Leipzig*). The lectures were organised in collaboration with the Donders Institute and the Radboud University Nijmegen by Petra van Alphen, Dan Dediu, Dieuwke de Goede, Mark Rijpkema and Nanjo Bogdanowicz.

FC Donders Lecture 2008

Sept. 23: Anthony T. Barker (*U. of Sheffield*).

'The interaction between electromagnetic fields and the human body: the birth and the development of Transcranial Magnetic Stimulation'.



