Social interaction influences the evolution of cognitive biases for language

Bill Thompson
bill@ling.ed.ac.uk
Language Evolution and Computation Research Unit, University of Edinburgh

Seán G. Roberts
sean.roberts@mpi.nl
Language and Cognition Department, Max Planck Institute for Psycholinguistics
www.ReplicatedTypo.com

Kenny Smith
kenny@ling.ed.ac.uk
Language Evolution and Computation Research Unit, University of Edinburgh

Introduction

We extend a computational Bayesian model of iterated learning with generations of agents learning multiple languages (Burkett & Griffiths, 2010). We show that the amount of linguistic diversity that emerges is affected by both the prior expectations of the learners and the social dynamics of the society in which they live.

Prior biases + Cultural evolution + Biological evolution + Communicative fitness = Linguistic diversity

A model of bilingual cultural transmission

There are two possible languages. Learners can use them in different proportions.

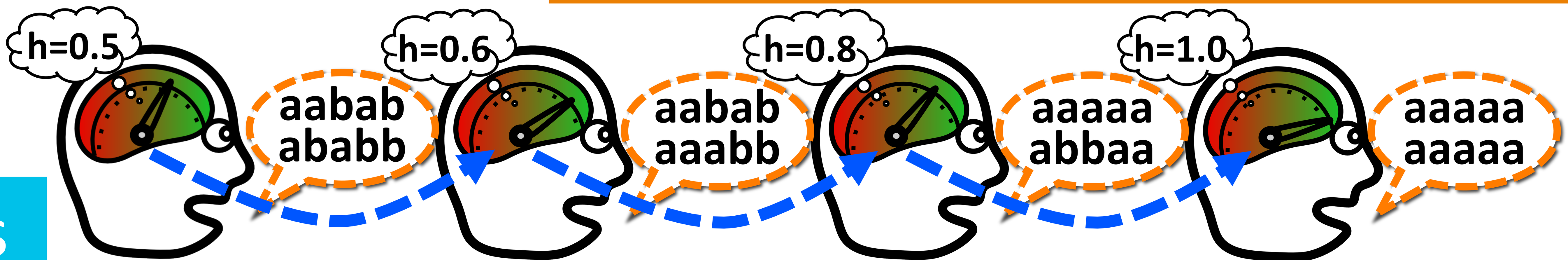
Learners observe a set of utterances d. Utterances can take one of two forms (u_1 or u_2) which are diagnostic of one of two languages l_0 & l_1, given some noise ε = 0.05:

$$P(u_i|l_i) = 1-\varepsilon \quad \text{and} \quad P(u_i|l_{j\neq i}) = \varepsilon$$

Each learner infers by Bayesian induction the frequencies of two possible languages, and induces a hypothesis, $h_0 = 1 - h_1 = P(l_0) \sim P(d|h_0)\,P(h_0)$

The likelihood is the product of probabilities for each function:

$$P(d|h) = \prod_{d_i} P(d_i|l_0)h + P(d_i|l_1)1-h$$



Prior Biases

Learners have two prior biases: One favouring the use of each language in a particular proportion (G_0), and one which controls the amount of variation they expect (α).

During inference, hypotheses h are drawn from a Dirichlet process prior with base distribution G_0 and concentration parameter **α**. G_0 specifies a distribution over the two possible language types. α regulates the influence of G_0 during inductive inference: as $\alpha \to \infty$, learners will induce hypotheses strongly determined by their prior preferences, so that $h \approx G_0$; as $\alpha \to 0$, h is determined largely by the learner's experiences. In the simulations below, $G_0 = (0.6, 0.4)$.

In our context we can interpret α as determining a learner's expectations about linguistic diversity. High α leads learners to expect a wide distribution of languages in the population. Low α leads learners to expect homogeneity: linguistic variation is discounted in favour of monolingual hypotheses.

Cultural evolution

Learners pass on their language culturally.

After a learner chooses a hypothesis, it produces utterances according to that hypothesis. These utterances become part of the input for the next generation of learners. We can track how the distribution of languages and priors changes over time. Previous work has shown that cultural evolution can amplify prior biases (see Smith & Thompson, 2012).

Biological evolution

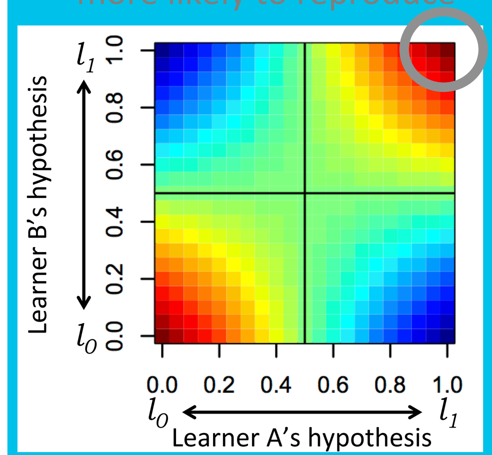
Learners pass on their prior biases (α) genetically.

A population of learners reproduce and pass on their prior biases to their children. Reproductive success depends on the agent's hypothesis and the fitness function (see below). The prior bias can mutate with probability $\mu = 0.01$, meaning that the distribution of priors evolves by natural selection.

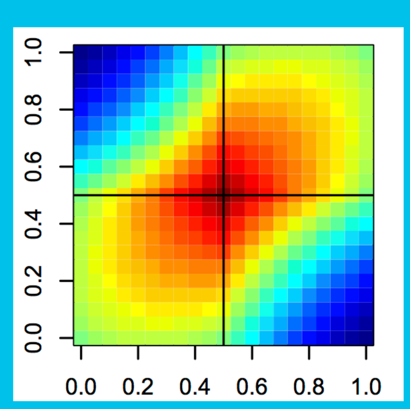
Communicative fitness

The probability of reproducing is based on social interaction. We test several metrics, including rewarding communicative success and rewarding bilingualism. The γ parameter controls the fitness payoff for knowing a second language.

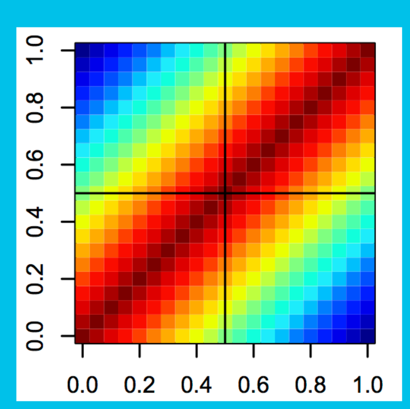
Monolingual
This metric rewards convergence to one particular language (monolingualism).
$\rho_m(h,h') = h \cdot h'$
This metric leads to low α (Smith & Thompson,2012).
Warmer colours = bigger payoff = more likely to reproduce



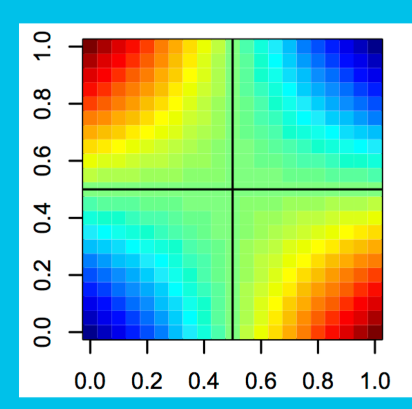
Bilingual
Agents that know both languages equally are fitter.
$\rho_b(h,h') = 2h \cdot h' \, (1-|h_0-0.5|) \times (1-|h'_0-0.5|)$
It pays to converge, but now on a single hypothesis. Balanced bilingualism is prestigious in many societies.



Parity
The maximum payoff is for learners who share the same hypothesis.
$\rho_p(h,h') = 1-|h_0 - h'_0|$
This removes any bias towards one particular language. Humans are sensitive to linguistic variation and match the behaviour of others (Gallois etal,2005;Smith&Wonnacott,2010)



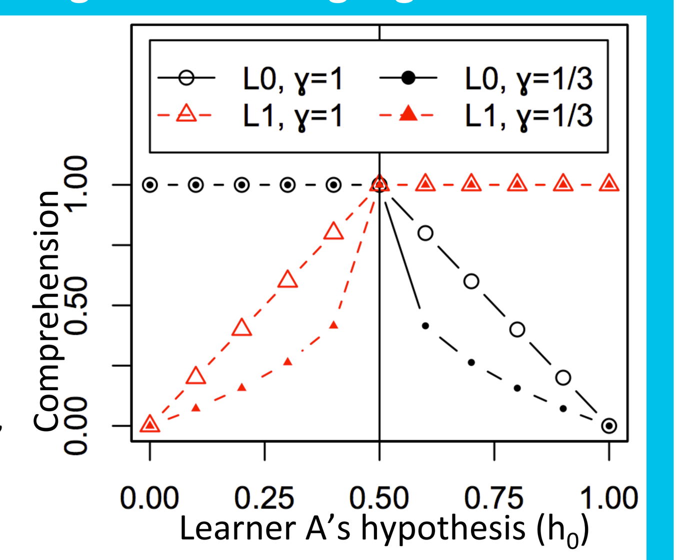
Linguistic Exogamy
Interactions with linguistically different individuals are rewarded
$\rho_{ex}(h,h') = 1-(h \cdot h')$
In human societies where marriage is restricted to different linguistic communities, bilingualism can persist (Hill, 1978; Jackson, 1983).



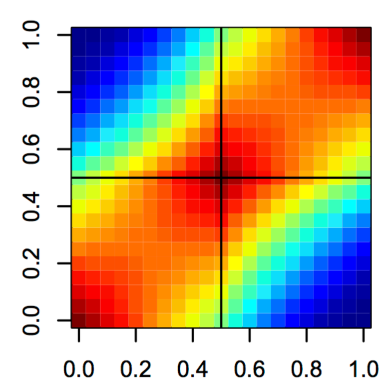
Dominant language
A speaker always understands its dominant language, and understands its non-dominant language in proportion to the balance of its hypothesis

$$\rho_d(h,h') = \begin{cases} 1 & \text{if } h > 0.5 \text{ and } h' > 0.5 \\ 1 & \text{if } h <= 0.5 \text{ and } h' <= 0.5 \\ |h-h'|^{\gamma} & \text{otherwise} \end{cases}$$

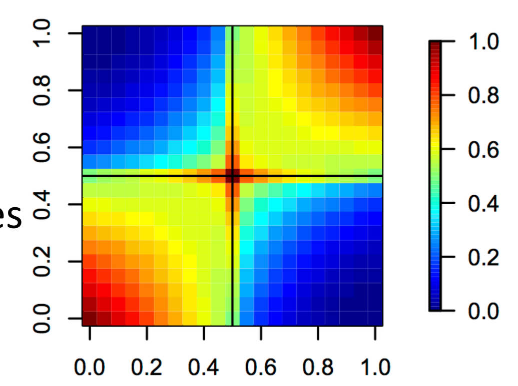
When two speakers have different dominant languages, the payoff is related to the difference between the hypotheses according to γ, which controls how much competence in an L2 an individual needs to get a good fitness payoff.



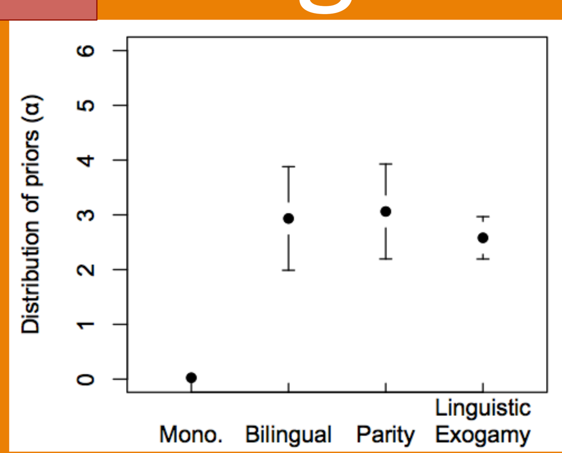
Fitness function for γ = 1. A small amount of L2 competence gives a good fitness payoff.

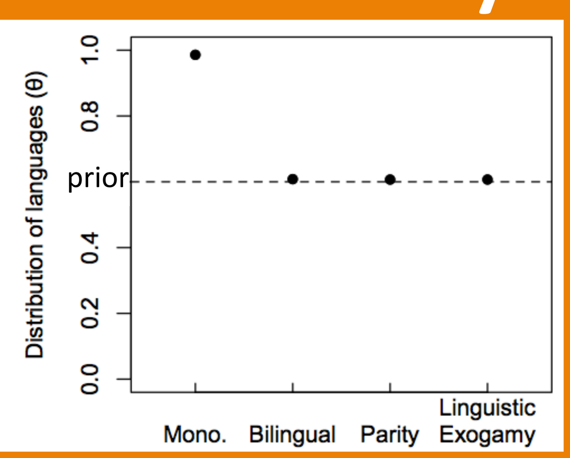


Fitness function for γ = 1/3. Only advanced competence gives a good overall fitness payoff.

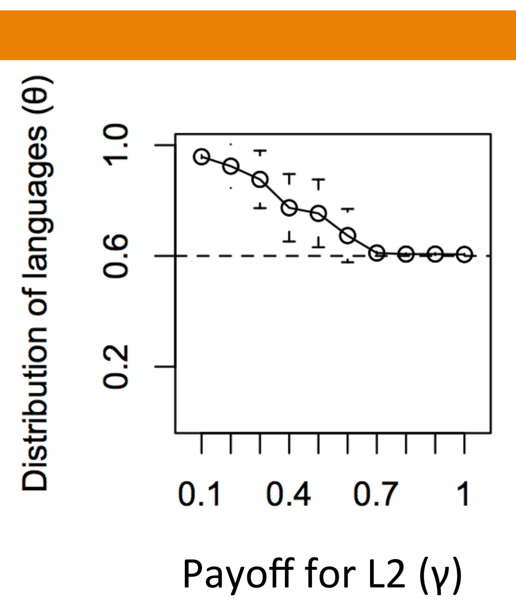


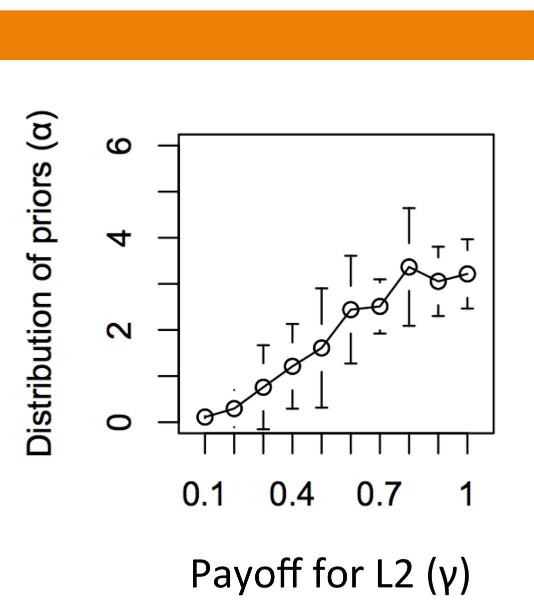
Linguistic diversity





In contrast with the 'monolingual' function, the alternative metrics result in bilingualism: the culture of these populations convergences to the prior over languages (G0), and is supported by learners who have evolved biases that favour linguistic diversity (high α). For the bilingual and exogamy metrics, this does not reflect the optimal fitness payoff.





As γ increases, there is a qualitative shift in the results of the simulations. With γ > 0.7 (competence in L2 is well rewarded), high α evolves (a 'bilingual' expectation) and the distribution of languages converges to the prior. However, with γ < 0.7 (competence in L2 is poorly rewarded), low α evolves and the distribution of languages is exaggerated (l_1 comes to dominate).

Conclusions

We used computational models to test a range of assumptions about how linguistic diversity is linked to individual biases and social interaction. The amount of diversity that emerges depends on both factors. Individual cognitive biases adapt to the context of social interaction. This shapes the effects of those biases on linguistic systems. In a range of situations, learners evolve to expect multiple languages in their input, even when acquiring competence in a second language is poorly rewarded.

References

Burkett, D., & Griffiths, T. (2010). Iterated learning of multiple languages from multiple teachers. In A. Smith, M. Schouwstra, B. de Boer, & K. Smith (Eds.), *The evolution of language: Proceedings of EvoLang 2010* (p. 58-65).

Smith, K., & Thompson, B. (2012). Iterated learning in populations: Learning and evolving expectations about linguistic homogeneity. In T. C. Phillips, M. Tamariz, E. A. Cartmill, & J. R. Hurford (Eds.), *The evolution of language: Proceedings of the 9th international conference (evolang9)* (p. 227-233). World Scientific.

Gallois, C., Ogay, T., & Giles, H. (2005). Communication accommodation theory: A look back and a look ahead. In Gudykunst, William B. *Theorizing About Intercultural Communication*. Thousand Oaks, CA: Sage.

Smith, K. and Wonnacott, E. (2010). Eliminating unpredictable variation through iterated learning. Cognition, 116(3):444–449.

Jackson, J. (1983). *The fish people: linguistic exogamy and Tukanoan identity in Northwest Amazonia.* Cambridge University Press.

Hill, J.H. (1978) Language contact systems and human adaptations. *J. Anthro. Research*, 34(1):1–26.