

# Project Sketch - Dablander & Görner

## Introduction

“One key challenge for cognitive science”, Kirby, Dowman, and Griffiths (2007, p. 5241) note, “is to explain the structure of human language.” One way of shedding light on this question is to consider how children acquire language. Traditional, bio-linguistic accounts (Pinker & Bloom, 1990), argue that most of our language knowledge is innate, formed by natural selection. In its extreme form, most of language learning is considered illusory (Smith, 2011). Recently, these viewpoints have been revised, and focus has been put on cultural factors. Since languages are culturally transmitted, they are potentially under selection for learnability. The basic argument is that languages adopt to become learnable by the agents, which leads to universal features like regularization and compositionality. The knowledge of language of one generation gets transmitted to the next generation and so on. The innate bias of the Agent - its cognitive constraints - form the language that we end up with.

## Iterated Learning Model

This process was formalized by Kirby in 1999 and termed "Iterated Learning Model". The three core features of this model (Ferdinand & Zuidema, 2008) are

1. Learning Algorithm
2. some form of information which is the input / output of the algorithm
3. structured transmission of information

Put crudely, this model formalizes games like “stille Post”. One agent starts with a word, say elk, and passes this word on to the next agent. Each agent represents a generation, and the word passed on represents a certain language. Not all information of the specific language (here the word "elk") gets passed onto the next generation (since in “stille Post” one has to whisper the word, thereby losing information). Each agent learns the language of the previous generation using some algorithm, and passes its variation of the language onto the next generation. In an early experiment using this paradigm, Bartlett had one participant draw an elk, and after some generations ended up with the drawing of a cat (Bartlett, 1932).

## Bayesian Rational Agents

While the standard account is to model agents as artificial neural networks (Kirby & Hurford, 2002), lately Bayesian models have come into fashion (Perfors, Tenenbaum, Griffiths, & Xu, 2011). They offer some unique benefits to modellers. The agent’s innate biases towards a certain language are explicitly encoded in the  $P(h)$ . The information that is passed on to the agent from the previous generation is  $P(d)$ . The agent computes

the likelihood of the utterance given a certain language,  $P(d|h)$ , for each language. Via Bayes' Rule

$$P(h|d) = \frac{P(h)P(d|h)}{\sum_h P(h)P(d|h)} \quad (1)$$

the agent computes the posterior distribution. Given the posterior distributions for each language, describing the probability of the data given each language, the agent picks one language using the specific learning algorithm. The agent then outputs some utterances using this language, serving as input for the next generation.<sup>1</sup>

Each learner sees some data, forms a hypothesis and then produces data according to that hypothesis, which serves as input for the next generation. This process across generations constitutes a markov chain, where the probability of a learner choosing a certain hypothesis, that is, acquire a certain language, depends only on the hypothesis chosen by the previous learner. We can construct a transition matrix which describes the probabilities that each hypothesis will lead to itself or any other hypothesis in the next generation (Ferdinand & Zuidema, 2008). This matrix holds all the possible trajectories of the iterated learning model. Once the model has converged, we can take the first eigenvector of the transition matrix, which yields the stationary distribution. The stationary distribution can be thought of as the expected distribution of languages that emerged due to cultural evolution (Kirby et al., 2007).

## Previous Results

Griffiths and Kalish (2005) demonstrated that iterated learning results in convergence to the prior distributions of hypothesis, i.e. the innate biases reflecting general cognitive capacities. Stated differently, the expected distribution of languages is proportional to the prior distribution, suggesting that language evolution is very much influenced by learning. It thus gives a detailed mechanistic account of how languages evolve through cultural transmission without relying on natural selection. However, by using not “proportional sampling” but MAP<sup>2</sup> as learning algorithm, Kirby et al. (2007) found that the prior is not a good predictor of the resulting distribution of languages. Instead, the relation between prior and posterior gets obscured by iterated learning; that is, the number of training examples - the cultural bottleneck - determines the posterior distribution. We can thus conclude that the type of learning algorithm is crucial in this simulation studies. When learners use a form of Gibbs sampling, the prior converges to the posterior. When MAP is used, the posterior is affected both by the prior and by the

---

<sup>1</sup>while assumptions of Bayesian rationality are controversial (Marcus & Davis, 2013), experiments in this domain (Kalish, Griffiths, & Lewandowsky, 2007) show that they might be reasonable.

<sup>2</sup>MAP is a point estimate which has the maximum a posteriori probability.

amount of information transmitted (Griffiths & Kalish, 2007).

In the studies mentioned above, each generation only consists of one individual. However, when learners learn from multiple individuals - a situation much closer resembling real life - different results emerge. Smith (2009) showed that the straightforward mapping from prior to posterior vanished when considering multiple agents per generation, and urges modellers to use a more realistic setting, possibly incorporating social structures.

## **Our approach**

We want to first replicate the results found when using one individual per generation using both Gibbs sampling and MAP as learning algorithm. More specifically, we want to investigate compositionality as a universal language feature. We will write our simulation in Python using the popular **NumPy** library, and possibly **PyMC** for Gibbs Sampling.

## References

- Bartlett, F. C. (1932). Remembering: a study in experimental and social psychology.
- Ferdinand, V. & Zuidema, W. (2008). Language adapting to the brain: a study of a bayesian iterated learning model.
- Griffiths, T. L. & Kalish, M. L. (2005). A bayesian view of language evolution by iterated learning. In *Proceedings of the 27th annual conference of the cognitive science society* (pp. 827–832).
- Griffiths, T. L. & Kalish, M. L. (2007). Language evolution by iterated learning with bayesian agents. *Cognitive Science*, 31(3), 441–480.
- Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007). Iterated learning: intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review*, 14(2), 288–294.
- Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104(12), 5241–5245.
- Kirby, S. & Hurford, J. R. (2002). The emergence of linguistic structure: an overview of the iterated learning model. In *Simulating the evolution of language* (pp. 121–147). Springer.
- Marcus, G. F. & Davis, E. (2013). How robust are probabilistic models of higher-level cognition? *Psychological science*, 24(12), 2351–2360.
- Perfors, A., Tenenbaum, J. B., Griffiths, T. L., & Xu, F. (2011). A tutorial introduction to bayesian models of cognitive development. *Cognition*, 120(3), 302–321.
- Smith, K. (2009). Iterated learning in populations of bayesian agents. In *Proceedings of the 31st annual conference of the cognitive science society* (pp. 697–702). Citeseer.