Effects of transmission perturbation in the cultural evolution of language

Thomas Brochhagen
University of Amsterdam

Michael Franke
University of Tübingen

Major factors in language evolution

- Efficient information transfer
- Learnability

Past research

Cognitive learning biases effect explanatory perturbations in language transmission (cf. Kirby et al 2007;2014)

But

We expect class of relevant transmission perturbations to be larger

(Iterated) Bayesian Learning

Learners consider the posterior probability of τ given a data sequence d of ⟨s, m⟩ pairs

$$P(\tau \mid d) \propto \underbrace{P(\tau)}_{\text{prior}} \quad \underbrace{P(d \mid \tau)}_{\prod_{\langle s,m\rangle \in d} P(m \mid s,\tau)}$$

With a parametrized posterior:

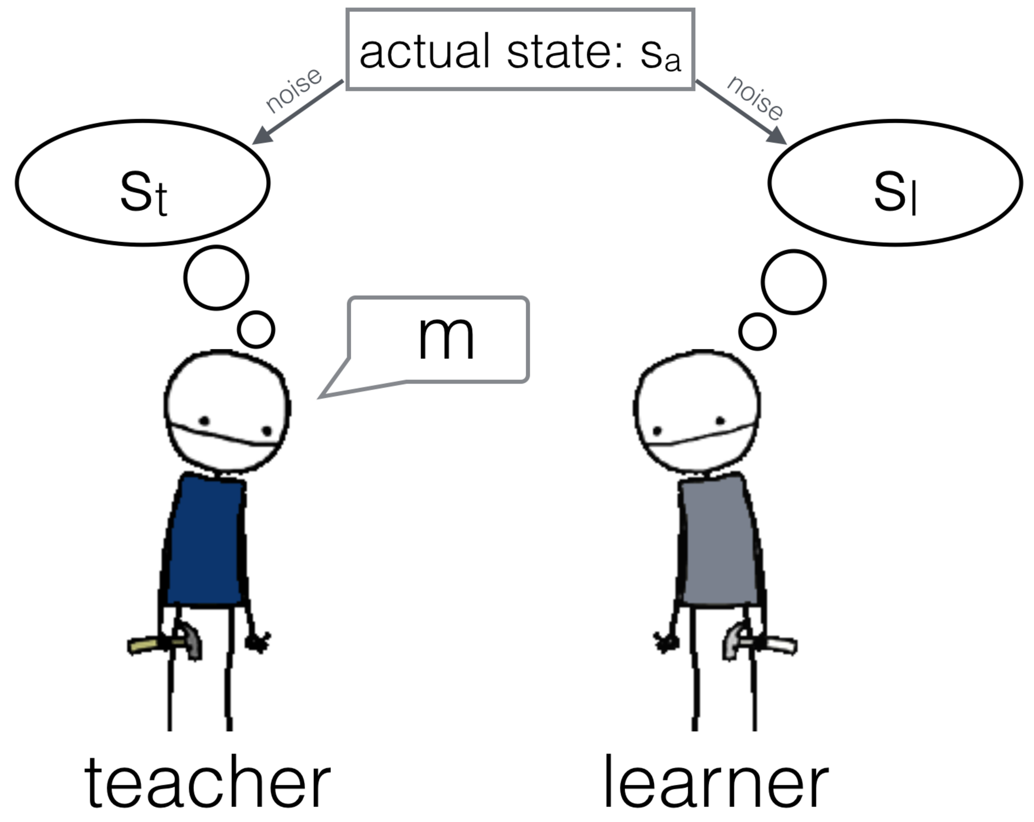
$$F(\tau \mid d) \propto P(\tau \mid d)^l, \quad l \geq 1$$

The probability that a learner acquires type j when learning from type i is then:

$$P(\tau_j \to \tau_i) \propto \sum_{d \in D_k} P(d \mid \tau_j) F(\tau_i \mid d)$$

Noise

(Iterated) learning with state-noise



The probability that s_a is the actual state when the learner observes s_l is

$$P_N(s_a \mid s_l) \propto P(s_a)\, P_N(s_l \mid s_a)$$

The probability that the teacher observes s_t when the learner observes s_l is

$$P_N(s_t \mid s_l) = \sum_{s_a} P(s_a \mid s_l) P_N(s_t \mid s_a)$$

The probability that type t produces a perceived sequence d_l is

$$P_N(d_l \mid \tau) = \prod_{\langle s_l,m\rangle \in d_l} \sum_{s_t} P_N(s_t \mid s_l)\, P(m \mid s_t, \tau)$$

The probability that a learner acquires type j when learning from type i is then:

$$P_N(\tau_j \to \tau_i) \propto \sum_{d \in D_k} P_N(d_l \mid \tau_j) F(\tau_i \mid d)$$

Conclusions

- Systematic noise in agents' perception of world states can give rise to stochastic perturbations that may influence and explain language change
- Noisy perception may mimic effects of cognitive biases and offers a neutral alternative to the study of language evolution under functional competition or differential learnability (cf. Reali & Griffiths 2009)

References: Brochhagen et al. (2016) in *Proceedings of CogSci* • Kirby et al. (2007) in *Proceedings of the National Academy of Sciences* • Kirby et al. (2014) in *Current Opinion in Neurobiology* • Reali & Griffiths (2009) in *Proceedings of the Royal Society B: Biological Sciences*

Case study I: Vagueness

- Two messages; hundred states
- Probability of perception of actual state as perceived state:

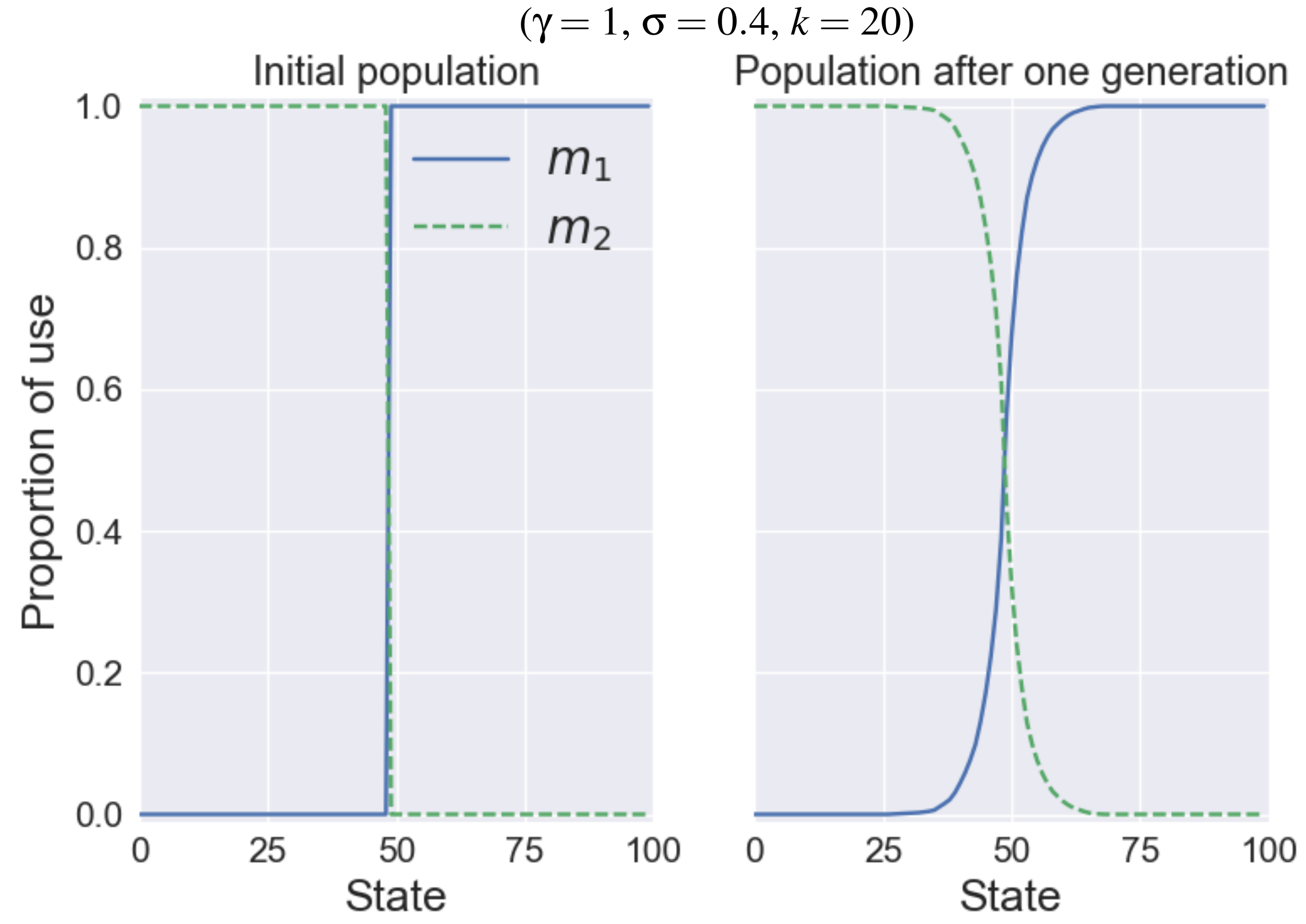
$$P_N(\cdot \mid s_a) \sim N(s_a, \sigma, 0, 99)$$

- Types:

$$\tau \in [0;99]$$

- Linguistic behavior:

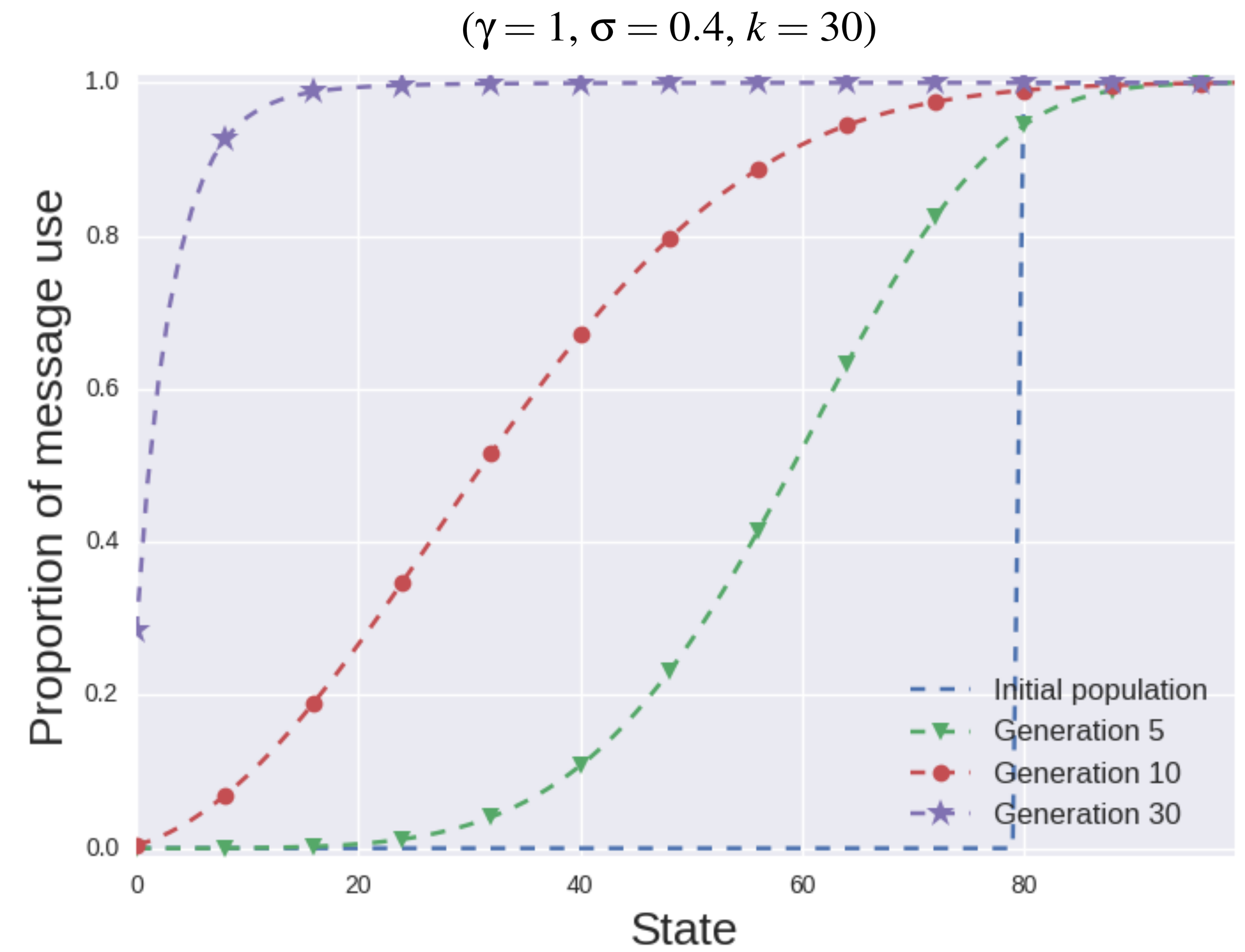
$$P(m_1 \mid s,\tau) = \delta_{s \geq \tau} = (1 - P(m_2 \mid s,\tau))$$



Case study II: Meaning deflation

- One message; hundred states
- Noise signature, behavior, and type space as in case study I
- We assume learners to take absence of observations into account:

$$P(\tau \mid d_l) \propto \text{Binom}(\text{successes} = k - |d_l|, \text{trials} = k, \text{succ.prob} = \sum_{i=0}^{\tau-1} P(s = i)) \prod_{s \in d_l} P(m \mid s,\tau)$$



Case study III: Scalar Expressions

- Previous research:
  - widespread adoption of non-upper bounded meaning in weak scalar expressions provided prior that favours them (Brochhagen et al. 2016)
- Target types:

	L_{bound}	L_{lack}
Literal	$\begin{array}{c\|cc} & m_{\text{some}} & m_{\text{all}} \\ \hline s_\forall & 0 & 1 \\ s_{\exists\neg\forall} & 1 & 0 \end{array}$	$\begin{array}{c\|cc} & m_{\text{some}} & m_{\text{all}} \\ \hline s_\forall & 0.5 & 0.5 \\ s_{\exists\neg\forall} & 1 & 0 \end{array}$
Pragmatic	$\begin{array}{c\|cc} & m_{\text{some}} & m_{\text{all}} \\ \hline s_\forall & 0 & 1 \\ s_{\exists\neg\forall} & 1 & 0 \end{array}$	$\begin{array}{c\|cc} & m_{\text{some}} & m_{\text{all}} \\ \hline s_\forall & 0 & 1 \\ s_{\exists\neg\forall} & 1 & 0 \end{array}$,

- Linguistic behavior:

$$P(m \mid s, \tau = \langle \text{lexicon}, \text{use} \rangle)$$

- Noise signature:

$$P(s_\forall \mid s_{\exists\neg\forall}) = \epsilon$$

$$P(s_{\exists\neg\forall} \mid s_\forall) = \delta$$