MODEL FITTING AND PREDICTION FOR LANGUAGE EVOLUTION

BILL THOMPSON, VANESSA FERDINAND

Language Evolution and Computation Research Unit,
The University of Edinburgh, Edinburgh EH8 9AD, UK
bill@ling.ed.ac.uk, vanferdi@gmail.com

As language evolution research becomes increasingly experimental, cognitive model fitting techniques are gaining attention (e.g. Culbertson et al., 2012). However, questions of language evolution must be seen in the light of cultural evolution. Here we note that this feature presents potentially unique difficulties in experimental design and model fitting. We consider several techniques that may help maximise confidence in the inferences researchers can draw about language learning biases and their consequences for population-level linguistic phenomena from experimental data. Two observations about culturally evolving systems motivate our methodologies: 1) learners will encounter a range of data, and 2) transmission may modulate the influence of learning biases. In order to get a good picture of the biases that shape language, it may be vital to know how learners respond to various kinds of data. Furthermore, characterising this range may allow us to make better predictions about crucial but unseeable experimental conditions, such as long-term transmission or structured populations. As such we ask: do certain approaches give a better picture of biases? and how can we best use this picture to generalise to populations?

We will illustrate these issues using data from Smith and Wonnacott (2010)'s regularisation experiments. Smith and Wonnacott's participants, organised into 10 Iterated Learning chains (ILCs) of 5 generations each, learned and reproduced the proportions of two competing plural markers in an artificial language. We first compare the inferences that can be drawn about Smith and Wonnacott's participants' regularisation biases had we only seen certain restricted subsets of their data, ignoring generation and focusing only on the marker proportions participants learned from[a]. Results suggest that only subsets that include multiple input proportions lead to conclusions reliably comparable to those based on the full dataset. In particular, organising participants into Iterated Learning chains mixes input proportions well.

---

[a]Inferences are drawn by fitting a beta-binomial learning model to participants' responses. Our procedure follows (Reali & Griffiths, 2009).

Based on the idea that the relation between cognitive biases and population level linguistic phenomena may be complicated by cultural evolution, we go on to examine a number of techniques for making predictions about unseen or unseeable experimental conditions of the kind central to cultural evolutionary theory (here Iterated Learning chains). Balancing input proportions across participants provides enough information to specify a learning model that can be used to reliably simulate cultural evolution. We consider various ways to do this, and benchmark each against the outcomes of Smith and Wonnacott's experimental Iterated Learning chains. In particular we test four methods for specifying the biases implemented in our simulations: 1) inferring from all experimental participants a single best-fit bias and using this uniformly in simulations; 2) inferring a separate best-fit bias for each participant and sampling this set in simulations, 3) inferring a separate best-fit bias for each participant, modelling the distribution of these biases, then sampling that distribution (or distributions) in simulations, and; 4) constructing a matrix of transition probabilities between input and output marker proportions directly from the data, and sampling these transition probabilities faithfully or with smoothing in simulations. Again, we test each method on various sparse subsets of Smith and Wonnacott's data (i.e. inferring biases from some subset of participants selected according to the input proportions they learned from). For these data, we achieve the closest fit to empirical chains with method 1) based on participant subsets in which all input proportions are evenly represented, though several of these approaches give good approximations.

While the merits of these methods will vary by case, we conclude that when designing learning experiments for language evolution research, it is helpful in some circumstances to divide participants among as many input conditions as is practicable: doing so can lead to a fuller picture of participants' biases as they would be expressed in a cultural system, and open up methods for generalising from experimental findings to population-level conclusions[b].

References

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*(3), 306–329.

Reali, F., & Griffiths, T. L. (2009). The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning. *Cognition*, *111*(3), 317–328.

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

[b]We thank Smith and Wonnacott for providing the data.