The evolution of frequency distributions: Relating regularization to inductive biases through iterated learning

Florencia Reali (floreali@berkeley.edu)
Thomas L. Griffiths (tom_griffiths@berkeley.edu)
Department of Psychology, 3210 Tolman Hall
Berkeley, CA 94720 USA

Abstract

The regularization of linguistic structures by learners has played a key role in arguments for strong innate constraints on language acquisition, and has important implications for language evolution. However, relating the inductive biases of learners to regularization has proven challenging. In this paper we explore how regular linguistic structures can emerge from language evolution by iterated learning, in which one person's linguistic output is used to generate the linguistic input provided to the next person. We use a model of iterated learning with Bayesian agents to show that this process can result in regularization when learners have the appropriate inductive biases. We then present two experiments demonstrating that simulating the process of language evolution in the laboratory can reveal biases towards regularization that might not otherwise be obvious, allowing weak biases to have strong effects. The results of these experiments suggest that people tend to regularize inconsistent word-meaning mappings.

Keywords: iterated learning; Bayesian models; frequency distributions; word learning; language acquisition;

Languages are passed from one generation of learners to the next via processes of cultural transmission. Such processes introduce linguistic variation, with the generalizations produced by each generation changing the prevalence of linguistic forms. A particular type of change occurs when forms with unpredictable or inconsistent variation become more regular over time. This process occurs in the creolization of pidgin and learning of sign languages from non-native speakers (e.g. Bickerton, 1981, see Hudson & Newport, 2005, for a review), and is often taken as evidence for innate language-specific constraints on language acquisition (e.g., Bickerton, 1981; DeGraff, 1999). This line of argument points toward the need to understand how the inductive biases of individual learners contribute to the regularization of inconsistent language forms. Identifying this relationship can provide insight into the constraints on the form of languages, and how words and grammars evolve over time. In this paper we begin to explore this question for the frequencies of linguistic variants.

Learning a language with any kind of probabilistic variation requires learning a probability distribution from observed frequencies. Over the last couple of decades, a number of studies have accumulated showing that learners are able to extract a variety of statistics from a wide range of linguistic input (see Gomez & Gerken, 2000; Saffran, 2003, for reviews). Recent work has explored how the *frequencies* of linguistic forms are learned. In this context, regularization corresponds to collapsing inconsistent variation towards a more deterministic rule. The empirical evidence as to whether this occurs with human language learners has been mixed. Hudson and Newport (2005) trained participants on artificial languages in which determiners occurred with nouns with varying probabilities. They found that adult participants produced utterances with probabilities proportional to their frequency in training (known as "probability matching"). However, they also found that children were much more likely to regularize, producing consistent patterns that were not the same as the training stimuli. Wonnacott and Newport (2005) extended these results using a similar artificial language, showing that when learners were tested on words different from those in the training stimuli, adults *did* regularize. However, other recent experiments seem to be at odds with this idea. For example, Vouloumanos (2008) examined how adults track the statistics of multiple-referent relations during word learning. Participants were trained on novel object-word pairs. Objects were associated with multiple words, which in turn were paired with multiple objects with varying probabilities. They were then presented with two objects while one of the words in was playing, and asked to select the object that went best with the word. The results indicated that participants tended to select responses in proportion to their frequencies, suggesting that people might probability match rather than regularize in learning multiple-referent relations.

The results outlined in the previous paragraph paint a mixed picture of the inductive biases involved in learning language from inconsistent input. In this paper, we take a novel approach to this problem. First, we outline a Bayesian model that can be used to make different kinds of inductive biases for frequency distributions explicit. We then use this model to characterize the consequences of a process of language evolution by iterated learning (Kirby, 2001), in which one learner's linguistic competence is acquired from observations of another learner's productions. This gives us a way to identify the conditions on the inductive biases of individual learners under which iterated learning results in regularization. We then simulate language evolution in the laboratory, using a variant on the task studied by Vouloumanos (2008) to show that while studying the responses of a single generation of participants does not reveal a bias towards regularization, this bias becomes extremely clear after a few generations of iterated learning. The results have implications for understanding both language evolution and language learning, revealing how weak biases can have a large effect on the languages spoken by a community, and how simulating language evolution in the laboratory can help to make these biases apparent.

A Bayesian model of frequency estimation

Our goal in studying the estimation of linguistic frequency distributions is to understand how the inductive biases of learners influence their behavior. To satisfy this goal, we need a formalism for describing learning that makes these inductive biases explicit. In this section, we outline how the frequency estimation problem can be solved using methods from Bayesian statistics. This allows us to identify how a rational learner with particular expectations about the nature of the frequency distributions in a language should behave, providing a basis for exploring the effects of these inductive biases on the evolution of frequency distributions and a method for inferring such biases from human behavior. Our focus will be on learning the relative frequencies of word-object associations. However, the models we develop apply to all problems that require learning probability distributions.

Assume that a learner is exposed to *N* occurrences of a referent (e.g., an object), which is paired with multiple competing linguistic variants with certain probability. We will use the example of estimating the relative frequency of two competing words, but our analysis generalizes naturally to larger numbers of variants, and to variants of different kinds. We will use x_1 to denote the frequency of word$_1$ (w_1) and $x_2 = N - x_1$ to denote the frequency of word$_2$ (w_2), and θ_1 and θ_2 to denote the corresponding estimates of the probabilities of these words. The learner is faced with the problem of inferring θ_1 and θ_2 from x_1 and x_2.

This estimation problem can be solved by applying Bayesian inference. The hypotheses being considered by the learner are all possible values of θ_1 (since θ_2 follows directly from this). The inductive biases of the learner are expressed in a *prior* probability distribution $p(\theta_1)$ over this set of hypotheses, indicating which hypotheses are considered more probable before seeing any data. The degrees of belief that the learner should assign to these hypotheses after seeing x_1 are the posterior probabilities $p(\theta_1|x_1)$ given by Bayes' rule

$$p(\theta_1|x_1) = \frac{P(x_1|\theta_1)p(\theta_1)}{\int P(x_1|\theta_1)p(\theta_1)\,d\theta_1} \tag{1}$$

where $P(x_1|\theta_1)$ is the *likelihood*, giving the probability of observing each value of x_1 for each value of θ_1.

For the case of two competing words, the likelihood $P(x_1|\theta_1)$ is defined by the Bernoulli distribution, with the probability of a particular sequence of *N* object-word pairings containing x_1 instances of w_1 is

$$P(x_1|\theta_1) = \theta_1^{x_1}(1-\theta_1)^{N-x_1} \tag{2}$$

where we assume that *N* is known to the learner. This likelihood is equivalent to the probability of a particular sequence of coin flips containing x_1 heads being generated by a coin which produces heads with probability θ_1.

Specifying the prior distribution $p(\theta_1)$ specifies the inductive biases of the learners, as it determines the conclusions that a learner will draw when given a particular value for x_1. We will assume that the frequency of w_1 and w_2 have a prior probability distribution given by a Beta distribution with parameters $\frac{\alpha}{2}$. This flexible prior corresponds to the distribution

$$p(\theta_1) = \text{Beta}(\frac{\alpha}{2}, \frac{\alpha}{2}) = \frac{\Gamma(\frac{\alpha}{2})}{\Gamma(\frac{\alpha}{2})\Gamma(\frac{\alpha}{2})}\theta_1^{\frac{\alpha}{2}-1}(\theta_2)^{\frac{\alpha}{2}-1} \tag{3}$$

where $\Gamma(\cdot)$ is the generalized factorial function (Boas, 1983).

The Beta distribution can take on different shapes depending on the values of α. As shown in Figure 1, when $\alpha/2 = 1$ the density function is simply a uniform distribution. When $\alpha/2 < 1$, the density function is U-shaped and when $\alpha/2 > 1$, it is a bell-shaped unimodal distribution. Thus, despite the apparent complexity of the formula, the Beta distribution captures prior biases that are intuitive from a psychological perspective. For example, when $\alpha/2 < 1$ the prior bias is such that the learner tends to assign high probability to one of two competing variants, consistent with regularization strategies. When $\alpha/2 > 1$, the learner tends to weight both competing variants equally, disfavoring regularization.

Substituting the likelihood from Equation 2 and the prior from Equation 3 into Equation 1 gives the posterior distribution $p(\theta_1|x_1,x_2)$. In this case, the posterior is also a Beta distribution, with parameters $x_1 + \frac{\alpha}{2}$ and $x_2 + \frac{\alpha}{2}$, due to the fact that the Bernoulli likelihood and Beta prior form a conjugate pair. The mean of this distribution is $\frac{x_1+\frac{\alpha}{2}}{N+\alpha}$, so estimates of θ_1 produced by a Bayesian learner will tend to be close to the empirical probability of w_1 in the data, $\frac{x_1}{N}$, for a wide range of values of α provided *N* is relatively large. Thus, even learners who have quite different inductive biases can be expected to produce similar estimates of θ_1, making it difficult to draw inferences about their inductive biases from these estimates.

Language evolution by iterated learning

Having considered how a single Bayesian learner should solve the frequency estimation problem, we can now explore what happens when a sequence of Bayesian learners each learn from data generated by the previous learner. In learning object-word relations, this corresponds to observing a set of objects being named, making an inference about the relative probabilities of the names, and then producing names for a set of objects which are observed by the next learner. More formally, we assume that each learner is provided with a value of x_1 produced by the previous learner, forms an estimate of θ_1 based on this value, and then generates a value of x_1 by sampling from $P(x_1|\theta_1)$, with the result being provided to the next learner. The key question is how the biases of the learners influence the outcome of language evolution via this process of iterated learning.

Griffiths and Kalish (2007) analyzed the consequences of iterated learning when learners are Bayesian agents. The first step in this analysis is recognizing that iterated learning defines a Markov chain, with the hypothesis selected by each learner depending only on the hypothesis selected by the previous learner. This means that it is possible to analyze the dynamics of this process by computing a *transition matrix*,

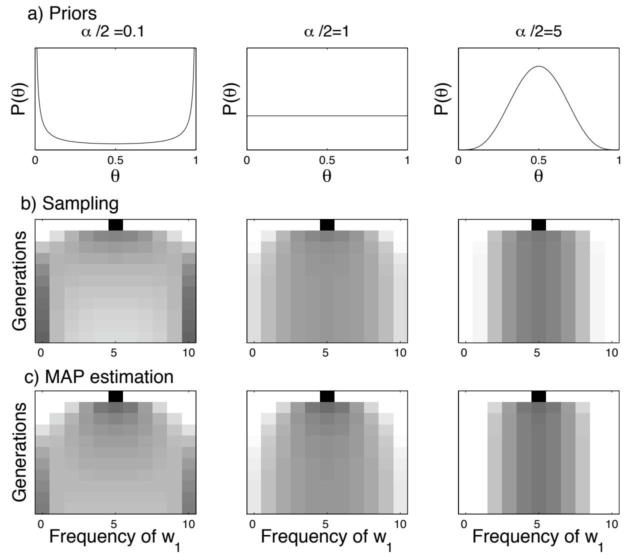

Figure 1: The effects of inductive biases on the evolution of frequencies. (a) Prior distributions on θ_1 for $\frac{\alpha}{2} = 0.1$ (left), $\frac{\alpha}{2} = 1$ (center), $\frac{\alpha}{2} = 5$ (right). Iterated learning by (b) sampling or (c) MAP estimation changes the probability distribution on the frequency of w_1 (horizontal axis) over several generations (vertical axis), but depends strongly on this prior. The frequency of w_1 was initialized at 5 from a total frequency of 10. White cells have zero probability, darker grey indicates higher probability.

indicating the probability of moving from one value of θ_1 to another or one value of x_1 to another across generations, and the asymptotic consequences by identifying the *stationary distribution* to which the Markov chain converges as the number of generations increases.

Further analysis of this Markov chain requires stating how the posterior distribution is actually translated into an estimate of θ_1. Griffiths and Kalish (2007) identifed two such estimation procedures: sampling a hypothesis from the posterior distribution, and choosing the hypothesis with the highest posterior probability. They demonstrated that when learners sample from the posterior, the stationary distribution of the Markov chain on hypotheses is the prior distribution. That is, as the number of generations increases, the probability of selecting a particular hypothesis converges to the prior probability of that hypothesis. In the case of frequency estimation, this means that we should expect that iterated learning with learners whose priors favor regularization (ie. with $\frac{\alpha}{2} < 1$) will ultimately produce strongly regularized languages.

It is typically more difficult to analyze the case where learners choose the hypothesis with highest posterior probability, known as the maximum *a posteriori* (MAP) hypothesis. However, in the case of frequency estimation the Markov chain defined by iterated learning is equivalent to a model that has been used in population genetics, the Wright-Fisher model of genetic drift with mutation (Ewens, 2004). This equivalence makes it possible to identify an approximate stationary distribution on θ_1, which is a Beta distribution with parameters $\frac{\alpha}{1+\frac{\alpha}{N}}$, where N is the total frequency (ie. 10 in the present example). A proof of this equivalence (which assumes that MAP estimation is performed in the natural parameter space of the Bernoulli distribution) is provided in Reali and Griffiths (2008). Unlike the case of sampling, frequencies do not converge to the prior distribution. However, the shape of the stationary distribution depends on the value of priors' parameter α. For example, it can be shown that for all values of $\alpha < \frac{N}{N-1}$, the stationary distribution is U-shaped.

The transition matrices associated with these two forms of estimation can also be computed. We will focus on the transition matrices for the values of x_1, as these values are easily observed in behavioral data. For the case of sampling, the probability that learner t generates a particular value of x_1 given the value generated by learner $t-1$ is given by

$$P(x_1^{(t)}|x_1^{(t-1)}) = \int P(x_1^{(t)}|\theta_1)p(\theta_1|x_1^{(t-1)})\,d\theta_1 \qquad (4)$$

where $P(x_1^{(t)}|\theta_1)$ is the likelihood from Equation 2 and $p(\theta_1|x_1^{(t-1)})$ is computed by applying Bayes' rule as in Equation 1. For the MAP case, the value of θ_1 produced as an estimate is deterministically related to $x_1^{(t-1)}$, so $P(x_1^{(t)}|x_1^{(t-1)})$ is given by Equation 2 with $\hat{\theta}_1 = \frac{x_1+\frac{\alpha}{2}}{N+\alpha}$ substituted for θ_1. These transition matrices can be used to compute the probability distribution $P(x_1^{(t)}|x_1^{(0)})$ as a function of the initial frequency of w_1, $x_1^{(0)}$, and the number of generations of iterated learning, t. The predictions of the sampling and MAP models are shown in Figure 1. Consistent with the analysis given above, the figure shows that when the prior distribution is bell-shaped, frequencies of linguistic variants converge over time to a distribution where the probability mass is concentrated around the mean. When the prior is U-shaped, the frequencies converge to a distribution where the probability mass is concentrated in the extremes of the distribution. Under these conditions, the most likely situation is that one variant becomes the vast majority in the population, while the other one becomes very infrequent, regardless of initial conditions. This situation can be interpreted as a regularization process.

The analyses presented in the last two sections support two conclusions. First, since the estimates of θ_1 produced by an individual learner will be only weakly affected by their prior, it can be hard to identify inductive biases by studying individual learners. Second, iterated learning can magnify these weak biases, resulting in rapid convergence to a regular language when learners have priors supporting regularization. These conclusions motivate the two experiments presented in the remainder of the paper. Experiment 1 demonstrates the difficulty of inferring the biases of learners by studying a single generation. Experiment 2 uses an iterated version of the same task to reveal that human learners favor regular languages, and to explore the consequences of this bias for language evolution by iterated learning.

Experiment 1: A single generation

The design of Experiment 1 was inspired by Vouloumanos (2008, Experiment 1). The experiment had a training phase where participants were exposed to novel word-object associations and a test phase assessing their knowledge of these associations. However, the design differs from Vouloumanos (2008) in that each word was associated with just one object, and the test trials consisted of a forced choice between words instead of objects.

Method

Participants Thirty undergraduates participated in the study in exchange for course credit.

Materials The materials used in Experiment 1 were the same used in Vouloumanos (2008). The auditory stimuli consisted of twelve words recorded by a native English female speaker. All words consisted of consonant-vowel-consonant syllables with consonants p, t, s, n, k, d, g, b, m, l and vowels æ, i, a, e, ∧ and u. Place of articulation was controlled both between and within words. Word pairs assigned to a common referent (object) were controlled so that they differed in the place of articulation, the vowels and letters they contained. The visual stimuli consisted of six out of the twelve three dimensional objects used in Vouloumanos (2008). The objects differed in color and shape and were animated to move horizontally as a cohesive unit. They were presented in short videos shown on a computer screen.

Design and procedure The experiment consisted of a training phase followed by a test phase. Participants were instructed that they would learn a novel language. No further information regarding the nature of the study was given in the instructions. During the training block participants were exposed to novel word-object associations. Each of the six objects were presented a total of 10 times, each time paired with one of two words (w_1 and w_2) with varying probabilities. The frequency with which each object occurred with w_1 and w_2 obeyed one of six different conditions. Conditions 0, 1, 2, 3, 4 and 5, corresponded to w_1 frequencies of 0, 1, 2, 3, 4 and 5, and w_2 frequencies of 10, 9, 8, 7, 6 and 5 respectively. For example, an object assigned to Condition 4 was presented 4 times with w_1 and 6 times with w_2 in the training phase. A unique pair of w_1 and w_2 was presented with a unique object. Therefore, the *overall* frequency of a word was determined by the frequency with which it appeared with its referent. Each of the six objects were randomly assigned to one of the six frequency conditions for every participant. The word pairs (w_1 and w_2) used to refer to each object were also randomized for every participant. On each trial, the object was presented for 3000 ms separated by 3000 ms, and the word was played concurrently with the visual stimuli. In addition to the auditory stimuli, the word was visually presented below the moving object.

The test block consisted of a forced choice selection task. Participants saw one object in the center of the screen and the two words associated with it where visually presented below the object image (bottom left and bottom right). Participants were instructed to select one of the two words pressing a key. The position of the word in the screen (left or right) was randomized across trials and participants. The six objects were presented 10 times each to match the number of presentations used in the training block. The order of training and test trials was randomized for every participant.

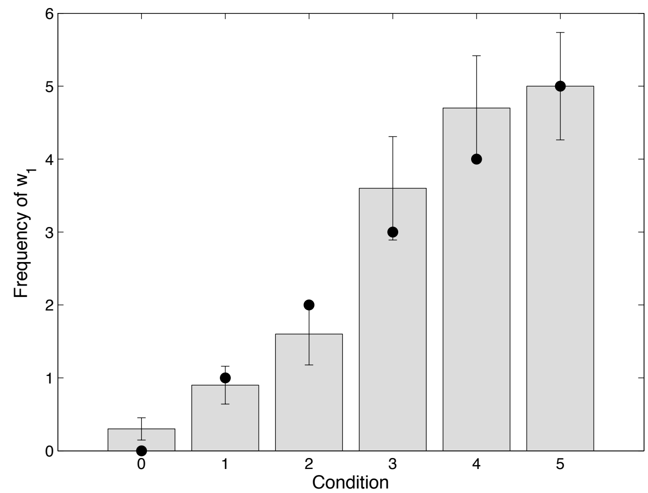

Figure 2: Results of Experiment 1, showing the mean frequency of w_1 selected by participants. Black dots correspond to w_1 frequency in the training stimuli and error bars indicate one standard error.

Results

There was a significant effect of w_1 frequency in the training stimuli on mean production of w_1 ($F(5,29) = 13.32, p < .0001$). In response to relative frequency values of 0, 1, 2, 3, 4, and 5 in the input, the mean number of w_1 in participants' productions were 0.3, 0.9, 1.6, 3.6, 4.7, and 5, respectively. Figure 2 compares the mean frequencies of w_1 produced by participants to the frequencies of w_1 in the training stimuli.

As shown in Figure 2, the mean frequency of w_1 in the productions was close to the corresponding frequencies in the training phase. However, this pattern of performance does not necessarily indicate that participants are probability matching rather than regularizing. The results displayed in Figure 2 are the group means and they could have resulted from averaging across individuals who each are using only one of the two competing words to name each object. To rule out this possibility, we examined the consistency of production among individual participants. We found that only 6 out of 30 participants regularized all of their productions.

The results of this experiment seem to suggest that people probability match when learning the probabilities with which words can be used to describe objects. These results are consistent with the conclusions of Vouloumanos (2008). However, the formal analyses presented above suggested that it may be difficult to detect a weak bias towards regularization

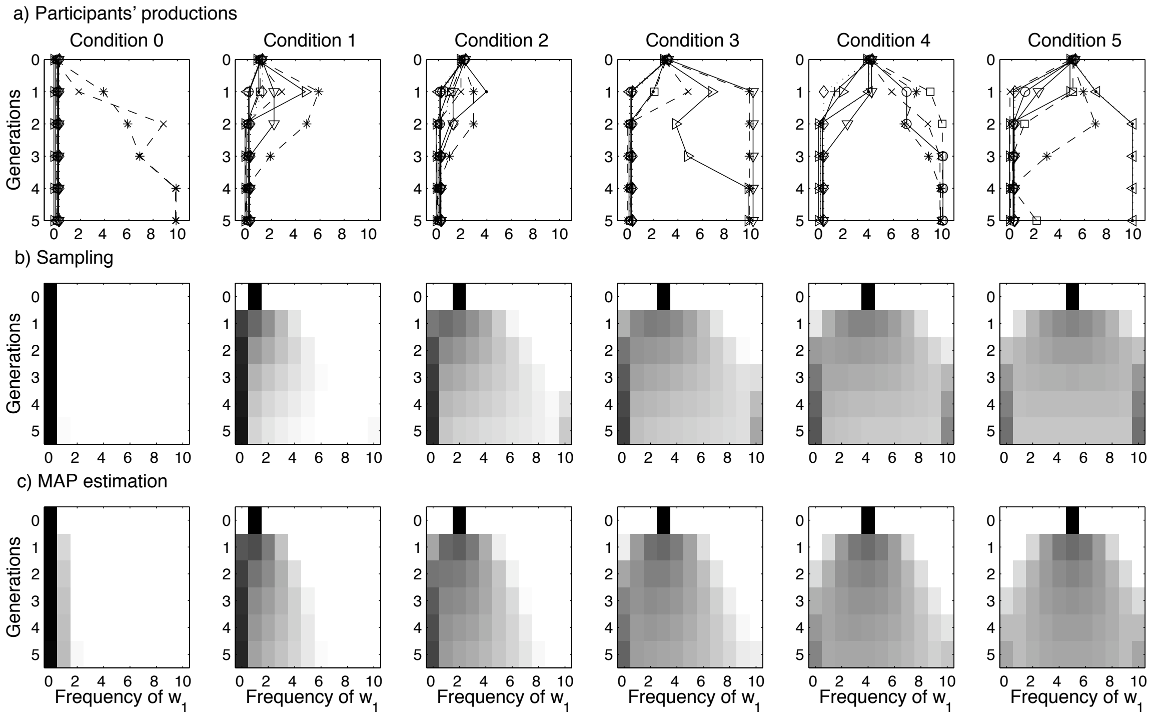

Figure 3: Results of Experiment 2. (a) Frequency of w_1 produced by participants (horizontal axis) per generation (vertical axis). Each panel corresponds to increasing values of the frequency of w_1 in the input to the first learner (right to left 0, 1 2, 3, 4, 5), and each line to one "family" of participants. Iterated learning with Bayesian agents using (b) sampling and (c) MAP estimation produce predictions in correspondence with these results. White cells have zero probability, darker grey indicates higher probability. The sampling model provides a better account of the participants' responses.

in a single generation of learners, even though such a bias might still have a significant effect on language evolution. Experiment 2 was designed to investigate the possibility that people have biases towards regularization that only emerge over several generations of iterated learning.

Experiment 2: Iterated learning

Method

Participants Fifty undergraduates participated for course credit. The participants formed five generations of learners in 10 "families". The responses of each generation of learners during the test phase were presented to the next generation of learners as the training stimuli.

Materials The materials used in Experiment 2 were the same as in Experiment 1.

Procedure For the 10 learners who formed the first generation of any family, the methods and procedure of the experiment were identical to Experiment 1. In subsequent generations, the method and procedure were the same, except that the frequency conditions in the training phase were determined by the productions of the previous participant within a family. That is, intergenerational transfer was implemented by letting the frequencies of w_1 (and w_2) produced by a single participant during the test phase be the frequencies of the training items for the participant in the next generation of that family. Participants were not made aware that their test responses would serve as training for later participants and intergenerational transfer was conducted without personal contact between participants.

Results

The results of Experiment 2 are shown in Figure 3. The top row shows participants' productions for each of the 10 families. The data is broken down across the six different initial conditions of relative frequency of w_1. Across all conditions, the frequencies of w_1 moved rapidly towards 0 and 10, reflecting a bias towards regularization. In fact, by the fourth generation, all productions were completely regular.

The sampling and MAP models introduced above were both fit to these data by maximum-likelihood estimation of the parameter α. The predictions of these models are shown in the middle and bottom rows of Figure 3. As can be seen from the figure, the models do a good job of capturing the dynamics of iterated learning. For the sampling model, the value of α that best fit the data was 0.05, giving a log-likelihood of -266, equivalent to a probability of 0.41 of correctly predicting the next value of w_1 from the previous one. For the MAP model, the value of α that best fit the data was 0.09, with a log-likelihood of -357, equivalent to a probability of

0.3 of correctly predicting the next value of w_1. These results suggest that the human data are better characterized in terms of learners sampling from their posterior distributions rather than MAP estimation.

Two aspects of the data are nicely captured by the model. First, as shown in the middle and bottom panels in Figure 3, the values of w_1 selected by learners in early iterations are close to the initial frequency of w_1. Thus, the model predicts responses that are consistent with probability matching when a single generation is considered. Second, the best fitting model is one where the prior distribution is U-shaped (see Figure 1, left panels). This means that the distribution over frequencies should converge to an equilibrium where one variant becomes the vast majority in the population, while the other one becomes very infrequent. Thus, the model predicts regularization of inconsistent language forms as a consequence of learners' inductive biases.

Discussion

Experiment 1 revealed that when participants were exposed to inconsistent word-meaning mappings, the frequencies determined by their responses were close to the frequencies present in training stimuli. This is consistent with the predictions of our Bayesian model. Moreover, the results are in accord with the pattern of responses reported by Vouloumanos (2008), which showed that participants were sensitive to fine-grained patterns of word-meaning mappings. The results of Experiment 2, however, revealed a trend toward regularization that was not obvious in a single generation. The distribution over competing words converged toward an equilibrium where one of the variants becomes the vast majority in the population. The dynamics of convergence again matched the predictions of our Bayesian model. This pattern of results indicates that weak regularization biases may have a strong effect on language change over time, and that iterated learning provides an effective method for revealing the inductive biases of human learners.

The results suggest that learners' inductive biases may favor regularization of inconsistent language forms. However, the question remains of how these inductive biases – represented in our model as a prior distribution – should be interpreted from a psychological viewpoint. The model's prior distribution should not necessarily be interpreted as the result of innate constraints specific to language. Rather, learning biases affecting the formation of linguistic representations could come from a number of domain-general constraints on learning, such as information-processing constraints, limitations on working memory; or the inductive bias associated with some kind of general-purpose learning algorithm. Another possibility is that the biases reflected by the model's prior distribution are not innately specified but the result of previous domain-specific experience.

Iterated learning models capitalize on the fundamental relation between language acquisition and language evolution. For example, certain types of language change may result from misalignments of the learner's and adult's hypothesis of the data (Pearl & Weinberg, 2007). This means that modeling language evolution may provide an effective way to empirically test hypotheses about language acquisition. Iterated learning offers a method to do this in the laboratory, connecting acquisition and evolution in a way that allows us to make contributions to understanding both of these processes. First, consistent with recent work on language evolution (Kirby, Dowman, & Griffiths, 2007), the experiments show how weak inductive biases can have strong effects in shaping linguistic distributions. Second, the results indicate that inductive biases can be hard to identify by testing individual learners, while they become evident in the context of language evolution. This suggests that a full understanding of the constraints on language acquisition might require the combination of multiple empirical approaches, including theoretical and empirical investigation of language evolution.

Acknowledgments. We thank Athena Vouloumanos for providing the materials used in the experiments, and Carla Hudson-Cam and Fei Xu for suggestions. We also thank Aaron Beppu, Matt Cammann, Jason Martin, Vlad Shut and Linsey Smith for assistance in running the experiments. This work was supported by grants BCS-0631518 and BCS-0704034 from the National Science Foundation.

References

Bickerton, D. (1981). *Roots of language*. Ann Harbor, MI: Karoma.

Boas, M. L. (1983). *Mathematical methods in the physical sciences* (2nd ed.). New York: Wiley.

DeGraff, M. (1999). Creolization, language change, and language acquisition: An epilogue. In M. DeGraff (Ed.), *Language creation and language change: Creolization, diachrony, and development* (p. 473-543). Cambridge: MIT Press.

Ewens, W. (2004). *Mathematical population genetics*. New York: Springer-Verlag.

Gomez, R., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178–186.

Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with bayesian agents. *Cognitive Science*, *31*, 441–480.

Hudson, C., & Newport, E. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*, 151–195.

Kirby, S. (2001). Spontaneous evolution of linguistic structure: An iterated learning model of the emergence of regularity and irregularity. *IEEE Journal of Evolutionary Computation*, *5*, 102–110.

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, *104*, 5241-5245.

Pearl, L., & Weinberg, A. (2007). Input filtering in syntactic acquisition: Answers from language change modeling. *Language learning and development*, *3*, 43–72.

Reali, F., & Griffiths, T. (2008). *Words as alleles: Equivalence of iterated learning and neutral models from population genetics.* (in preparation)

Saffran, J. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*, 110–114.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*. (in press)

Wonnacott, E., & Newport, E. (2005). Novelty and regularization: The effect of novel instances on rule formation. In *BUCLD 29: Proceedings of the 29th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press.