Cognition 111 (2009) 317–328

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

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

Florencia Reali, Thomas L. Griffiths *

University of California, Berkeley, Department of Psychology, 3210 Tolman Hall # 1650, Berkeley, CA 94720-1650, United States

ARTICLE INFO

Article history:
Received 18 September 2008
Revised 13 January 2009
Accepted 15 February 2009

Keywords:
Iterated learning
Bayesian models
Frequency distributions
Word learning
Language acquisition

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 behavior in laboratory tasks can be challenging without a formal model. 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 three 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, and that even a weak bias towards regularization can allow regular languages to be produced via language evolution by iterated learning.

1. Introduction

Languages are passed from one learner to the next via processes of cultural transmission. Such processes introduce linguistic variation, with the generalizations produced by each learner changing the prevalence of linguistic forms. A particular type of change occurs when components of language with unpredictable or inconsistent variation lose their unpredictability and become more regular over time. This process of *regularization* has come to play a prominent role in discussions of the role of innate constraints on language acquisition in linguistics and cognitive science (e.g., Bickerton, 1981; Pinker, 1994).

An example of regularization appears in the creolization of Pidgin languages and certain forms of learning of sign languages (e.g. Bickerton, 1981; Siegel, 2007, see Hudson Kam & Newport, 2005, for a review). Pidgin languages typically emerge when speakers of mutually intelligible languages come together and need to communicate. Speakers then create a new communication system based on the superstrate language, that is, the language that predominates in the region. *Creoles* are more stable forms of language that originate as pidgin and which are learned by the children of a community as their native language. Pidgin or early Creole languages contain variability that is not typical of natively acquired languages (Birdsong, 1999, Johnson, Shenkman, Newport, & Medin, 1996). For example, speakers are inconsistent in their use of morphological markers or word order. Importantly, unlike the kind of variation present in native speech, which is typically predictable and shared by all speakers, the variation in pidgin languages is largely unpredictable – as is typical of second language productions. For example, in early Hawaiian Creole the particular word order used by individual speakers was influenced by the word order used in their native language. In later stages of Creole, however, language forms typically lose this unpredictability and become more regular. Further evidence that learners exposed to incon-

* Corresponding author. Tel.: +1 (510) 642 7134; fax: +1 (510) 642 5293.
E-mail address: tom_griffiths@berkeley.edu (T.L. Griffiths).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.02.012

sistent use of grammatical forms tend to regularize their input comes from learning during acquisition of sign language from inconsistent teachers (Singleton & Newport, 2004) and from the emergence of regular systems in the creation of new sign languages (Senghas & Coppola, 2001; see Hudson Kam & Newport, 2005, for a review).

Another example of regularization occurs in situations of language contact where unpredictable variability is introduced to a language. Consider the case of word order changes observed in the transition from Old to Modern English. Scandinavian influence is thought to have introduced the verb-object order to English, resulting in a temporarily mixed system composed of verb-object and object-verb word orders in Old English (Kroch & Taylor, 1997). Over time, however, verb-object gradually replaced object-verb word order, resulting in the regular system observed in Modern English (for a review see Pearl & Weinberg, 2007).

The tendency of learners to regularize inconsistent language forms has often been taken as evidence for innate language-specific constraints on language acquisition (e.g., Bickerton, 1981, 1999). For example, according to the Language Bioprogram Hypothesis (Bickerton, 1981), when children are exposed to reduced communication systems such as pidgin languages, they introduce universal properties of natural languages by drawing on their innate knowledge of natural language structure. Recent studies, however, have seriously challenged some of the fundamental tenets of the Language Bioprogram Hypothesis. The emergence of creole appears to be less abrupt than previously assumed and seems to depend on stabilized forms of pidgin spoke by adults as a second language (Siegel, 2007). This points toward the need to understand how the inductive biases of individual learners – those factors that constrain their inferences from limited linguistic data – contribute to the regularization of unpredictable variability.[1] Identifying this relationship can provide insight into why languages take the forms they do, and how words and grammars evolve over time. In this paper we begin to explore this question for the case of estimating 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, 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. In one study, Hudson Kam and Newport (2005) trained participants on artificial languages in which determiners occurred with nouns with varying probabilities. They found that children regularize the unpredictability in the input, producing consistent patterns that were not the same as the training stimuli. They also found that adult participants produced utterances with probabilities proportional to their frequency in training, a response referred as *probability matching*.

Subsequent studies by Hudson Kam and Newport (in press) showed that, for adult participants, regularization depends on the form and complexity of the inconsistency in the input. For example, when two variant forms in the artificial grammar were used in free alternation, that is, the determiner being either present or absent in a sentence, the most frequent form was not regularized. However, when many different determiners were used and one form was much more frequently and consistently used than the others, adults *did* regularize that most consistent form. In a different study, Wonnacott and Newport (2005) used a similar artificial language to show that when adults learners were tested on words different from those in the training stimuli, adults regularized. Taken together, these results suggest that the level of complexity in the probabilistic input might influence whether learners adopt a regularization strategy rather than probability matching.

Another recent study on word learning provides further insights into the learning biases operating during learning from inconsistent input. 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 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 studies outlined in the previous paragraphs suggest that language learners regularize under some circumstances, and probability match under others. However, identifying the inductive biases influencing frequency estimation can be challenging. Without a formal model that translates the inductive biases of learners into explicit predictions about behavior, it can be hard to determine what evidence a particular empirical result provides about those biases. For example, rather than a simple dichotomy between probability matching and regularization, we might imagine that biases towards regularization vary continuously in their strength, with different expectations, task demands, and processing limitations determining the strength of the bias in a given context. A formal model of the effects of inductive biases on frequency estimation would provide a way to make this distinction, and its predictions could be used to design experiments that test whether a given task results in probability matching or just a weaker bias towards regularization.

In this paper, we use a Bayesian model to make explicit the inductive biases that operate during frequency estimation of language forms. This model allows us to characterize the consequences of cultural transmission by *iterated learning* (Kirby, 2001) – the process by which one learner's

[1] While claims about innate constraints on language learning are clearly making statements about the inductive biases of learners, our use of the term should not be interpreted as only reflecting such constraints. Inductive biases relevant to language acquisition could come from a variety of domain-general factors, including innate constraints, a point that we return to in the General Discussion.