Information and Computation 209 (2011) 536–547

Contents lists available at ScienceDirect

Information and Computation

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

Efficiency of the symmetry bias in grammar acquisition

Ryuichi Matoba [a,*], Makoto Nakamura [b], Satoshi Tojo [b]

[a] *Department of Electronics and Computer Engineering, Toyama National College of Technology, 1-2, Neriya, Ebie, Imizu, Toyama 933-0293, Japan*
[b] *School of Information Science, Japan Advanced Institute of Science and Technology, 1-1, Asahidai, Nomi, Ishikawa 923-1292, Japan*

ARTICLE INFO

Article history:
Available online 18 November 2010

Keywords:
Symmetry bias
Abduction
Iterated Learning Model
Compositionality
Grammar acquisition

ABSTRACT

It is well known that the symmetry bias greatly accelerates vocabulary learning. In particular, the bias helps infants to connect objects with their names easily. However, grammar learning is another important aspect of language acquisition. In this study, we propose that the symmetry bias also helps to acquire grammar rules faster. We employ the Iterated Learning Model, and revise it to include the symmetry bias. The result of the simulations shows that infants could abduce the meanings from unrecognized utterances using the symmetry bias, and acquire compositional grammar from a reduced amount of learning data.

1. Introduction

Infants have an amazing capacity to learn their vocabularies. In fact, infants can acquire new words very rapidly, e.g., 7–15 words a day over the age of 18 months, and can also learn a word's meaning after just a single exposure, through fast mapping [1–5]. The difficulty in lexical acquisition includes mapping between a word and a meaning from an infinite range of possible meanings, as is well known as 'gavagai problem' [6]. Some cognitive biases are working for lexical acquisition. Infants overcome the difficult mapping through fast mapping and slow mapping. The process of fast mapping forms an approximate representation of a word's meaning after just a single exposure on the basis of heuristics [7], or cognitive biases such as whole object bias, mutual exclusivity bias, and symmetry bias. Carey [1] explained that the approximate word meanings were first established by fast mapping, and then were embodied as actual words in infants' memory through slow mapping. Therefore, infants under 17 months old only learn lexical items slowly, but this knowledge is fragile and those lexical items are prone to being forgotten. On the other hand, infants over 18 months old come to acquire new words firmly, and lexical misapplication subsides [8].

Among these various biases, the *symmetry bias* is said to be saliently effective for lexical learning [9–14]. The bias says, if infants are taught that an object P has a lexical label Q, then they apply the label Q to the object P. For example, an infant who has learned the name of an object, which is called `apple`, can not only answer the name of the object indicated correctly, but can also pick out an apple from a basket filled with other fruits. The latter action is considered as a human-specific skill, and many experiments have shown that other animals cannot map an object to its label [15–17]. Although the relation between symmetry bias and lexical acquisition has been reported so far, its relation to grammatical construction has not been mentioned yet. In this paper, we verify the efficacy of the symmetry bias not only in lexical acquisition, but also in grammar acquisition.

Our study is based on the Iterated Learning Model (ILM, hereafter) by Kirby [18]; in each generation, an infant can acquire grammar in his mind given sample sentences from his mother, and a grown infant becomes the next mother to speak to a new-born baby with his/her grammar. As a result, infants can develop more compositional grammar through the generations. Note that the model focuses on the grammar change over multiple generations, not on that in one generation. In ILM's setting, a sentence is considered to be uttered in an actual situation; thus, an infant can guess what a sentence implies. Therefore,

* Corresponding author.
E-mail addresses: rmatoba@nc-toyama.ac.jp (R. Matoba), mnakamur@jaist.ac.jp (M. Nakamura), tojo@jaist.ac.jp (S. Tojo).

0890-5401/$ - see front matter
doi:10.1016/j.ic.2010.11.018

Fig. 1. Holistic rule.

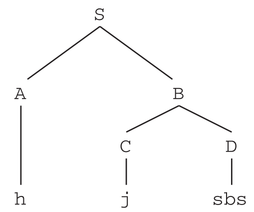

Fig. 2. Compositional rule.

a mother's utterance is always paired with its meaning. Here, the utterance is regarded as an instance of E-language, while the infant guesses its grammar structure, that is, I-language [19–22]. Our objective in this paper is to embed the symmetry bias into this combination of utterance and meaning, to show its efficacy.

This paper is organized as follows: in Section 2 we explain Kirby's ILM [18], and in Section 3 we revise it to include the symmetry bias. Section 4 presents the details of our experimental model, and gives specific experiment designs. We analyze our experimental results in Section 5, and conclude and discuss our results in Section 6.

2. Iterated Learning Model

ILM is a framework for investigating the cultural evolution of linguistic structure. In the experiment by Kirby [18], a parent is a speaker agent and her infant is a learner agent. The key issue of this model is the learning bottleneck, that is, the infant acquires sufficiently versatile grammar in spite of the limited amount of sentence examples from the parent; in other words, this bottleneck is the very cause of grammar generalization. The infant tries to guess sentence structure, as utterances are always paired with their meanings, which are intrinsically compositional. This process is iterated generation by generation, and finally in a certain generation an agent would acquire a compact, limited set of grammar rules.

2.1. Utterance rule of Kirby's model

According to Kirby's model, we present a signal–meaning pair as follows:

$$S/hit(john, ball) \rightarrow \text{hjsbs} \tag{1}$$

where the meaning, that is the speaker's intention, is represented by a Predicate–Argument Structure (PAS, hereafter) *hit*(*john*, *ball*) and the signal is the utterance `hjsbs`; the symbol '*S*' stands for the category *S*entence. The following rules can also generate the same utterance.

$$S/hit(x, ball) \rightarrow \text{h}\,N/x\,\text{sbs} \tag{2}$$

$$N/john \rightarrow \text{j} \tag{3}$$

The variable x in (2) can be substituted for an arbitrary element of category N. A rule without variables, i.e., the whole signal indicates the whole meaning of a sentence as in Formula (1) is called a *holistic rule* (Fig. 1), while a rule with variables as in Formula (2) is called a *compositional rule* (Fig. 2). Also, a rule which corresponds to a word, as in Formula (3) is called a *lexical rule*. Here, we review the definitions of these rules.

Definition 1 (*Compositional rule, Holistic rule, and Lexical rule*)**.**
Compositional rule: a grammar rule including non-terminal symbols for categories.
Holistic rule: a grammar rule consisting of terminal constants.
Lexical rule: a rule consisting of a monadic terminal constant.