Language Sciences 38 (2013) 53–58

Contents lists available at SciVerse ScienceDirect

Language Sciences

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

Beyond tier-based bigrams: an artificial grammar learning study

Hahn Koo [a], Young-il Oh [b,*]

[a] Department of Linguistics and Language Development, San José State University, United States
[b] Department of English, Seoul National University of Science and Technology, South Korea

A R T I C L E I N F O

Article history:
Received 10 August 2012
Received in revised form 14 January 2013
Accepted 15 January 2013
Available online 17 February 2013

Keywords:
Nonadjacent dependency
Phonology acquisition
Artificial grammar learning
Statistical learning

A B S T R A C T

Some of recently proposed phonotactic learners are tier-based bigram learners that restrict their hypothesis space to patterns between two segments that are adjacent at the tier level. This assumption is understandable considering that typologically frequent nonadjacent sound patterns are predominantly those that hold between two tier-adjacent segments. However, it is not clear whether the assumption is psychologically justified, i.e., whether speakers are indeed exclusively attentive to patterns between two tier-adjacent segments when it comes to learning nonadjacent sound patterns. In general, many recent studies suggest that learnable sound patterns are not limited to typologically observed sound patterns. Specifically, Koo and Callahan (2012) argue that adult speakers in laboratory settings have no trouble learning artificial patterns that cannot be explained by tier-based bigram learners. In this paper, we replicate their results in a more carefully controlled setting and argue that the assumption of tier-based bigram learning must be relaxed in order to properly explain human performance.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Some of recent proposals on how phonological dependencies between nonadjacent segments can be learned are what Heinz (2010) calls tier-based bigram learners (Newport and Aslin, 2004; Hayes and Wilson, 2008; Goldsmith and Riggle, 2012). Segments in each word are projected onto relevant phonological tiers (Goldsmith, 1976) and the learner scans pairs of adjacent units, i.e., bigrams, within each tier to look for statistical regularities. This effectively limits the hypothesis space of the learner to dependencies between two segments that are adjacent at the tier level. Many nonadjacent sound patterns often discussed in the literature or considered typologically frequent are characterized as local dependencies between two sounds that are next to each other at the tier level. The following quote by McCarthy and Prince (1996, p. 1) reflects this well: "a rule may fix on one specified element and examine a structurally adjacent element and no other."

Given the prevalence of phonological rules that "count up to two" (McCarthy and Prince, 1996, p. 1), the hypothesis space of tier-based bigram learners is arguably a reasonable approximation of typologically observed sound patterns. However, it is not clear whether the hypothesis space reasonably approximates what speakers can actually learn. The types of patterns that speakers can learn are often studied experimentally using the artificial grammar learning paradigm (Reber, 1967). In a typical experiment using the paradigm, subjects are first exposed to words that exemplify some artificial pattern and then perform a task that allows the investigator to tell whether they have learned the pattern from the exposure. Recent studies using the paradigm strongly suggest that learnability of a sound pattern is not necessarily predictable from typological data. Some

* Corresponding author. Address: Department of English, Seoul National University of Science and Technology, 232 Gongneung-ro, Nowon-gu, Seoul 139-743, South Korea. Tel.: +82 2 970 6283; fax: +82 2 973 8349.
E-mail address: youngoh@seoultech.ac.kr (Y.-i. Oh).

0388-0001/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.langsci.2013.01.004

studies report successful learning of unattested or rare sound patterns in laboratory settings (e.g., Seidl and Buckley, 2005), sometimes as well as typologically frequent sound patterns (e.g., Koo and Cole, 2006).

Dependencies between temporally nonadjacent units in general are considered harder to learn than those dependencies between temporally adjacent units. Saffran et al. (1996) report that adult speakers can detect statistical regularities between adjacent syllables upon hearing a stream of CV syllables. However, Newport and Aslin (2004) report that adult speakers seem unable to detect the same regularities between two syllables separated by an intervening syllable. Gomez (2002) reports a study in which adults and 18-month olds could learn dependencies between nonadjacent words upon hearing three-word sequences, but only when the variability of the intervening word was high enough.

Unlike dependencies between nonadjacent syllables or words, many artificial grammar learning studies report successful learning of dependencies between two nonadjacent phonemes. Despite being segmentally nonadjacent, however, the co-dependent phonemes in the vast majority of such studies are structurally adjacent to each other within some phonological tier: the consonant tier (e.g., Bonatti et al., 2005), the vowel tier (e.g., Pycha et al., 2003), the sibilant tier (e.g., Finley, 2012), etc. These studies show that the hypothesis space of the tier-based bigram learners should be large enough to successfully learn the nonadjacent dependencies that human subjects could learn in those studies. But results of these studies also suggest that the hypothesis space must be further constrained as well since there were artificial dependencies between two tier-adjacent phonemes that subjects could not learn. For example, adult subjects in Pycha et al.'s (2003) study failed to learn an arbitrary dependency between stem final and suffix vowels (e.g., the suffix vowel must be /ɛ/ instead of /ʌ/ if the stem final vowel is {/i/, /æ/, /ʊ/}).

Koo and Callahan (2012), however, argue that tier-based bigram learners are too restrictive to be a model of how speakers learn nonadjacent sound patterns, based on two artificial grammar learning experiments. In the first half of each experiment, adult English speakers were familiarized with "study" words of $C_1V_1.C_2V_2.C_3V_3$ structure that exemplify dependencies either between C_1 and C_2 (Experiment 1) or between C_1 and C_3 (Experiment 2). In the latter half, they heard a set of "test" words and decided for each word whether they had heard it in the first half or not. Some of the words were novel but "legal" words that respected the dependencies, while other words were novel and "illegal" words that violated the dependencies. Subjects in both experiments mistook the legal words as familiar words more often than they mistook the illegal words.

The authors attribute the results to learning of phonotactic dependencies. For example, legal words in their Experiment 2 contained the same C_1–C_3 subsequence also found in study words, while illegal words contained novel C_1–C_3 subsequence not found in any of the study words. Due to the presence of familiar C_1–C_3 subsequences that embody the dependencies in legal words, subjects must have found legal words more similar to study words than illegal words are. As a result, subjects confused legal words with study words more often than they confused illegal words.

However, there seems to be an alternative way to explain the results. Analysis of the materials used in the experiments suggests that tracking the Hamming distance (Hamming, 1950), i.e., number of word positions occupied by different segments, between test words and study words could also lead subjects to prefer legal words to illegal words. More specifically, the Hamming distance of legal words (e.g., /salemu/) from their closest study words (e.g., /salemi/) in the experiments turns out to be 1.0, while the corresponding distance of illegal words (e.g., /sameli/) is 2.0. That is, subjects could have compared each test word with study words one segment at a time independent of other segments in the word and still found legal words to be more similar to study words than illegal words are.

This alternative possibility weakens Koo and Callahan's argument against psychological plausibility of tier-based bigram learners. Accordingly, we ran an artificial grammar learning experiment similar to Experiment 2 of Koo and Callahan's study with the difference in the Hamming distance controlled for and tested whether the results of Koo and Callahan can be replicated. We discuss the details of the experiment in Sections 2 and 3 below.

2. Methods

We tested learnability of dependencies between C_1 and C_3 embedded in words of the form $C_1V_1.C_2V_2C_3$ in two versions of an artificial grammar learning experiment, henceforth Experiments A and B. The basic idea is similar to that of Koo and Callahan (2012). Subjects were first familiarized with study words and then heard a novel set of test words—either legal or illegal words—and rated how familiar they sounded on a five-point scale. We interpreted a higher mean familiar rating to legal words than to illegal words as evidence of learning. The two versions of the experiment functioned as control experiments for each other: what counts as legal words in Experiment A counts as illegal words in Experiment B, and vice versa. This was to ensure that any preference towards legal words in an experiment is due to learning rather than some inherent bias towards legal words over illegal words.

2.1. Subjects

Fifteen students at Seoul National University of Science and Technology participated in each version of the experiment with a total of 30 subjects for the current study (15 × 2 = 30). All subjects were native speakers of Korean without any hearing loss or speech impediment. No subject participated in both Experiment A and Experiment B.

Download English Version:

https://daneshyari.com/en/article/1103210

Download Persian Version:

https://daneshyari.com/article/1103210

Daneshyari.com