Paradigm uniformity in the lab: prior bias, learned preference, or L1 transfer?

Adam Albright (MIT) and Youngah Do (Georgetown University)

Participants in artificial grammar (AG) learning experiments frequently favor uniform paradigms, even when they are trained on languages that show consistent stem alternations (Pater and Tessier 2003; Wilson 2006; White 2014). For example, Albright and Do (2013) trained participants on singular∼plural noun paradigms in which 100% of stop-final items alternated in voicing or continuancy, and yet 27% of participant responses failed to apply alternations to unseen items in the testing phase, favoring non-alternating paradigms instead. This preference mirrors errors in language acquisition (Kazazis 1969; Do 2013) and diachronic change (Malkiel 1968; Schindler 1974), raising the intriguing possibility that artificial grammar learning experiments can be used to investigate experimentally the forces that drive paradigm leveling in language change. A prior (innate) paradigm uniformity preference is not the only explanation for non-alternation errors in the lab, however. It is also possible that participants favor non-alternation because of experience with their native language (L1 transfer) or with non-alternating filler items (a learned preference; Albright 2005). We present here the results of an AG experiment in which we manipulate the amounts of evidence available from non-alternating fillers, and also compare how participants generalize to segments that do or do not alternate in their native language. As in previous studies, we find significant numbers of non-alternating responses, even for stem types that showed consistent alternations in training. Furthermore, the rate of non-alternation is not increased by exposure to more non-alternating fillers, or by non-alternation in L1. The only factor that systematically affects non-alternation is greater exposure to alternating items; we model these results with a MaxEnt learning model, incorporating a prior paradigm uniformity bias (OO-Faithfulness constraints).

To test the influence of alternating and non-alternating items in the acquisition of alternations, we devised a set of 5 artificial languages, independently varying the number of each: 4, 8 or 12 alternating items with 8 non-alternators, and 4, 8 or 12 nonalternators with 8 alternators. If learners start with a prior bias for non-alternation (e.g., high-ranking OO-Faith constraints), we predict that increasing the number of alternating items should increase the probability of generalizing alternations, while increasing the number of non-alternating items should have no effect, since learners already expect non-alternation. Conversely, if learners simply match the relative proportion of alternating items in the data, generalization of alternations should increase with more alternators, and decrease with more non-alternators. In the artifical languages, stop-final stems showed voicing alternations (seɪp ∼ seɪbi, bɹik ∼ bɹigi), while fricative- and nasal-final stems showed no alternations (dɹun ∼ dɹuni, kluf ∼ klufi). American English has extremely productive voicing alternations for coronals (flapping: *weigh*[t] ∼ *weigh*[ɾ]*y*) and limited voicing alternations for fricatives (*hou*[s]*e* ∼ *hou*[z]*es*), but no systematic voicing of /p/ or /k/. Thus, it is also possible to detect L1 transfer by examining generalization to /t/-final stems: if participants are using English rankings, then they should apply voicing alternations to /t/ (and perhaps even /s/) at higher rates than /p/ or /k/.

Alternating and non-alternating items were embedded in an implicit learning task, in which participants were instructed to pay attention to the plural suffix (-i after consonants, -nu after vowels); the total number of training items was held constant across all languages. The test items included novel stop-final items (trained /p/, /k/; withheld /t/) and fricative-final items, and participants were forced to choose between unvoiced (non-alternating) and voiced (alternating) plurals. 250 adult native English speakers participated on Amazon Mechanical Turk, and responses were analyzed using mixed effects logistic regression. The results show that voicing was generalized more often to trained p, k than to t (β=.15, t=4.10). This is unexpected if participant preferences are due to L1 transfer, since English actually favors voicing alternations for coronals. Furthermore, the number of alternating responses increased systematically with the number of alternating training items (Fig. 1), but did not decrease systematically with the number of non-alternating training items (not shown). This is expected if learners are employing a prior bias for non-alternation, which can be suppressed with data from alternating items, but need not be reinforced with data from non-alternating items. It is not expected, however, if learners are simply mimicking the proportion of alternating items in the training data. Thus, the results support a model in which artificial language learners employ a prior bias for non-alternation, mirroring the OO-FAITH $\gg$ MARKEDNESS bias that has independently been posited for L1 acquisition (McCarthy 1998; Hayes 2004; Do 2013).

We model these results using a regularized Maximum Entropy model of weighted constraints (Goldwater and Johnson 2003; Jäger 2007). We implement the paradigm uniformity bias by including OO-FAITH constraints in the grammar, with a prior weight above that of MARKEDNESS. Since the model initially favors non-alternation and learning is error-driven, the grammar changes only in response to alternating training items, and not to non-alternators; this captures a key finding of the experiment. Learning alternations consists (in part) in promoting MARKEDNESS constraints that favor alternations, such as *V[obstruent]V, *V[voiceless stop]V, and *V[k]V, so that their cumulative weight exceeds that of OO-FAITH. In the MaxEnt framework, training on p∼b and k∼g results in promotion of all relevant markedness constraints, so that they 'share the credit' for explaining alternations. A prediction of this approach, illustrated in Fig. 2 is that alternations should be generalized most to trained alternations, where specific constraints such as *VkV also support them, and less to untrained alternations, where the relevant specific constraints have not been promoted (e.g., *VtV). This captures the second experimental finding: participants only partially generalized alternations from /p/, /k/ to /t/, in spite of the fact that English favors voicing alternations for coronals.

The results here are novel, in that few prior AG experiments on phonological alternations have been designed to rule out L1 transfer as the source of an observed bias. (In fact, most studies cited above do not even discuss the possibility of L1 transfer.) Furthermore, the modeling results show that an independently successful model of L1 grammar learning can capture important properties of how humans learn and generalize phonological patterns in an artificial grammar setting. Of course, this alone does not prove that participants in AG experiments are solving the task in exactly the same way that they learn L1, but it does contribute to a growing body of work showing that AG experiments are a useful tool in probing the mechanisms of L1 language learning.

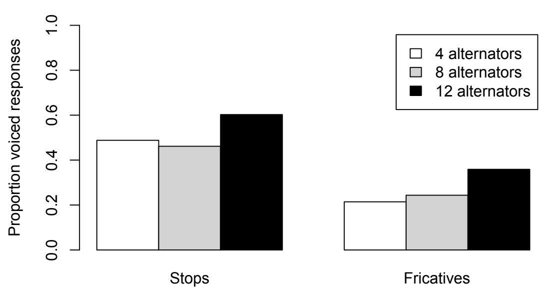

Figure 1: Voiced responses increase with number of alternating training items

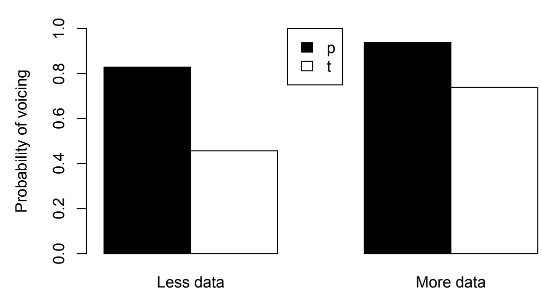

Figure 2: Model predictions: partial generalization to unseen /t/