Proc Ling Soc Amer Volume 1, Article 11:1-15
doi:http://dx.doi.org/10.3765/plsa.v1i0.3709

Learning consequences of derived-environment effects

Adam J. Chong*

Abstract. In constraint-based phonological models, it is hypothesized that learning phonotactics first should facilitate the learning of phonological alternations. In this paper, we investigate whether alternation learning is impeded if static phonotactic generalizations and dynamic generalizations about alternations mismatch as in derived-environment patterns. English speakers were trained on one of two artificial languages, one in which static and dynamic generalizations match (Across-the-board), the other where they did not (Derived-environment). In both languages, there was an alternation that turned [ti] and [di] to [ʧi] and [ʤi] respectively across a morpheme boundary. In the Across-the-board language, the constraint motivating this (*Ti) was true across-the-board, whereas words with such sequences within stems were attested in the Derived-environment language. Results indicate that alternation learning in both languages was comparable. Interestingly, learners in the Across-the-board language failed to infer the *Ti constraint despite never hearing words with such sequences in training. Overall, our results suggests that alternation learning is not hindered by a static phonotactic mismatch in this type of experimental paradigm and that learners do not readily extend a generalization about legal heteromorphemic sequences to analogous sequences within a morpheme.

Keywords. Artificial language learning; phonological acquisition; derived-environment effects; phonotactics; alternations; duplication problem

1. Introduction. It has been observed that phonological alternations that arise due to morpheme concatenation often reflect static phonotactic restrictions that hold true of the lexicon. That is, phonological alternations in multimorphemic words serve to ensure that a surface string conforms to the phonotactic restrictions that hold of monomorphemic words. This observation is often called the ‘duplication problem’ (Kenstowicz & Kisseberth 1977). For example, in Navajo (Fountain 1998; Kari 1976; McDonough 1991; McDonough 2003; Sapir & Hoijer 1967), sibilants in a stem must agree for the feature [anterior] (/ʧ/ and /ʒ/ are [-anterior]; /ts/ and /s/ are [+anterior]). Thus one only observes harmonic roots in Navajo; disharmonic roots are unattested (data from Martin 2011).

(1) /ʧ’oʒ/ (*ts’oʒ)
worm
‘worm’

(2) /ʦ’ózí/ (*ʧ’ozí)
slender
‘slender’

* I would like to thank Robert Daland, Bruce Hayes, Sharon Peperkamp, Megha Sundara and Kie Zuraw, as well as audiences at UCLA Phonology Seminar, USC PhonLunch, AMP 2015 and LSA 2016 for discussion and feedback. I am also very grateful to Audrey Kim for help with initial piloting of the experiments, Eleanor Glewwe for recording stimuli, as well as Michael Becker and Daniel Szeredi for assistance with Experigen. This research was funded by a UCLA Dissertation Year Fellowship. Author: Adam J. Chong, University of California, Los Angeles (ajchong@ucla.edu).

This same constraint also triggers phonological alternations across a morpheme boundary as in (3) and (4) (from Martin 2011). The prefix /-s-/ agrees in anteriority with the sibilant in the root, alternating in (3) to harmonize with /ʒ/ in the root. Thus in Navajo, what is true of the lexicon also holds of phonological alternations.

(3) /ji-s-lééʒ/ → [ji-ʃ-lééʒ]
4SG.SBJ-PERF-paint
'it was painted'

(4) /ji-s-tiz/ → [ji-s-tiz]
4SG.SBJ-PERF-spin
'it was spun'

This state of affairs however is not always the case. One class of phonological patterns in which static phonotactic generalizations of the lexicon and generalizations about alternations are argued to mismatch are derived-environment effects (a.k.a. non-derived environment blocking; Burzio 2011; Kiparsky 1973; Kiparsky 1993; Wolf 2008, *a.o.*; see also Paster (2013) for other cases). A well-known example comes from Korean (e.g. Cho 2001; Iverson & Wheeler 1988). In Korean, /t/ and /t^h/ palatalizes to /c/ and /c^h/ preceding the high front vocoids /i/ and /j/, resulting in the alternations in (5-6).

(5) /mat-i/ → [maci]
eldest-NOM
'eldest'

(6) /pat-i/ → [paci]
field-NOM
'field'

However, this alternation does not seem to serve any phonotactic aim since /ti/ and /t^hi/ sequences are attested within stems (7-8). Thus, there is a mismatch between the stem generalization which allows /ti/ and /t^hi/ sequences and the generalization that drives the alternation (*ti).

(7) /mati/ → [mati]
knot
'knot, field'

(8) /titi-ta/ → [titita]
'tread-COP'
'to tread'

In constraint-based phonological frameworks such as Optimality Theory (OT; Prince & Smolensky 1993), both static phonotactic generalizations as well as generalizations driving alternations are often captured by a single constraint. If phonotactic learning occurs first before alternation learning, then the latter is facilitated by the initial learning of phonotactics (Hayes 2004; Hayes & Wilson 2008; Prince & Tesar 2004; Tesar & Prince 2007; Jarosz 2006). This predicts, however, that when there is a mismatch between phonotactics and alternations (as in derived-environment effects), learners should find alternations more difficult to learn since they might initially arrive at an incorrect phonotactic generalization. For example, a learner of Korean might incorrectly learn to accept /ti/ sequences across-the-board when learning from a morphologically unparsed signal, only to have to learn later that such sequences alternate heteromorphemically.

Overall, experimental studies on phonological learning have focused on either the learning of phonotactic knowledge alone (e.g. Skoruppa & Peperkamp 2011; Linzen & Gallagher 2014) or the learning of phonological alternations alone (e.g. Wilson 2006; Cristià et al. 2013; White 2014). These studies, however, have largely ignored the relationship between these two levels of generalization. One study, Pater and Tessier (2005), stands out in this regard. Pater and Tessier (2005) examined whether American English listeners learned a phonological alternation more easily when the alternation was motivated by a phonotactic constraint that is active in the English lexicon. Participants were trained on one of two languages. In Language 1, participants learned a language with [t]-epenthesis that occurred when a word would have surfaced with a final lax vowel. This alternation is motivated by the fact that, in English, lax vowels cannot occur word-finally. In Language 2, participants learned the same alternation, but in this case [t]-epenthesis occurred following front vowels word-finally. The alternation in Language 2 was not motivated by phonotactics in English. They predicted that if learners utilize their existing knowledge of the phonotactics of English, then the alternation should be easier to learn in Language 1 than in Language 2. Pater and Tessier found, in accordance with their expectations, that participants learned the alternation better in the language with English phonotactic support (Language 1) than the one without (Language 2). However, as the authors themselves point out, not only is the alternation in Language 2 not consistent with English phonotactics, it is also typologically unnatural. Given that we know that unnatural patterns are at the very least dispreferred by learners (e.g. Hayes et al. 2009; Becker, Ketrez & Nevins 2011; Hayes & White 2013), it is possible that the poorer performance in Language 2 could be explained by this alone.

The current studies are designed to address this confound by testing learners in an artificial grammar learning paradigm (e.g. Wilson 2006) on two different languages that had the same, equally natural, phonological alternation (modeled on Korean palatalization), but that differed only in whether or not the constraint motivating the alternation was true stem-internally. That is, in one language, the static phonotactic generalization and the dynamic one motivating the alternation match, whereas in the other, they mismatch. This controlled manipulation ensures that any difference we see in alternation learning can be attributed to the match/mismatch status of these two sets of generalizations. If phonotactics aids in alternation learning, we expect participants to learn the alternation better when these generalizations match than when they mismatch.

2. Experiment 1. In Experiment 1, learners were trained on two artificial languages that had the same alternation but differed on whether the phonotactic constraint motivating this alternation was true stem-internally. In a departure from previous work in artificial grammar learning, participants were tested on two types of phonological knowledge across two test phases: a blick test (Halle 1978) designed to test static phonotactic knowledge, and a wug test (Berko 1958) that examined what they learned about phonological alternations. If learners in the Across-the-board language infer a phonotactic gap, they should endorse words with [ti] and [di] sequences less often than comparable fillers. Derived-environment language learners should endorse both of these equally since these sequences appeared in the training data. Given the mismatching phonotactic information in the Derived-environment language, we also expect that the phonological alternation in this language should be more difficult to learn than in the Across-the-board language.

2.1. ARTIFICIAL LANGUAGES. Two artificial languages were created with alternations modeled after /t/-palatalization in Korean described above. The languages were constructed using the consonants {p, t, ʧ, b, d, ʤ} and the vowels {a, i, u}. Singular words were of three types: $C_1V_1C_2V_2C_F$ (e.g. *bitap*), $C_1V_1C_2V_2$ (e.g. *padu*), and $C_1V_1C_F$ (e.g. *pib*). Primary stress was always on the initial syllable. Plurals in both languages involved the suffixation of /-i/ (e.g. *bapi*). In both languages, stem-final [t] and [d] became [ʧ] and [ʤ] respectively preceding the suffix [-i]. Stem-final [p, b, ʧ, ʤ] remained unchanged. When stems were vowel-final, the plural allomorph consisted of an epenthetic [ʔ] (e.g. *padu'i* [paduʔi]).

In the Derived-environment language, 20 $C_1V_1C_2V_2C_F$, 20 $C_1V_1C_2V_2$, and 18 $C_1V_1C_F$ stems were created. In order to maximize the phonotactic CV gap in the Across-the-board language, we sought to increase the independent occurrence of [t, d] (with [a, u]) and the vowel [i] (with [p, b]). This was done by having more of such CV sequences than other CV sequences (7 each of [pi] and [bi], and 6 each of [ta], [tu], [da] and [du]). All other CVs occurred equally frequently (5 each) within stems. The frequencies of CV sequences in both languages is summarized in Table 1. CV frequencies including forms with the plural suffix attached were not controlled for. In the stem-final C_F position, 21 singular stems ended in non-alternating consonants [p, b, ʧ, ʤ] and 13 ended with alternating [t, d]. In order to guard against the urge to make a product-oriented generalization (i.e. plurals should have [ʧ] or [ʤ] regardless of the base; e.g. Bybee 2001), only 4 stems ended in [ʧ, ʤ]. This meant that less than a third of plurals would have surface [ʧi] or [ʤi]. Crucially, in the Derived-environment language, the sequences [ti] and [di] were found in stems, resulting in a mismatch between the dynamic generalization ([t-i]/[d-i]→[ʧ-i]/[ʤ-i]) and the static phonotactic generalization ([ti] and [di] are legal). In total there were 58 training items in this language. An overview of training stimuli in both languages with regards to alternations is shown in Table 2.

CV type	Example singular stem	Across-the-board	Derived-environment
{t, d}i	tip	--	10
{t, d}{a, u}	taʧup	24	24
{p, b}i	ʤubi	14	14
{ʧ, ʤ}{i}	ʤit	10	10
{p, b, ʧ, ʤ}{a, u}	dubad	11	13

Table 1. No. of CVs in each language

Stem-final consonant	Singular		Plural	No. Across-the-board	No. Derived-environment	%
V	padu	→	paduʔi	18	20	~35
p, b:	dupib	→	dupibi	19	21	~37
ʧ, ʤ:	pubiʤ	→	pubiʤi	4	4	~7
t, d:	dat	→	daʧi	11	13	~21
			Total	52	58	

Table 2. Overview of plural formation in each language

The Across-the-board language consisted of mostly the same stems as in the Derived-environment language. The key difference was that words with [ti] and [di] were excluded. The leftover syllables were recombined to make 4 words to ensure that each CV type was presented equally frequently in both languages. Due to this, however, there were only 52 training singulars in the Across-the-board language. It is important to note that although the learning data in both languages differed slightly, participants in both languages heard the exact same frequency of filler CVs (non-[ti] and [di]) as shown in Table 1, and the relative proportion of alternating vs.

non-alternating final consonants was also kept as similar as possible given the differences described above as shown in Table 2. Also, while the number of stems that occurred with particular stem-final consonants was different, the proportion of stems with each consonant in the language was kept as similar as possible across both languages. Apart from these small differences, the crucial difference between the two languages was that in the Across-the-board language, learners never heard [ti] or [di] sequences in training; learners in the Derived-environment language did. This means that any differences we might see in performance between both language groups can be attributed to the absence or presence of such sequences in training.

The same test words were used in both language groups. Novel words used in both test phases were formed in a similar way as in training. In the blick test, 36 novel words were created: 12 [ti/di] items and 24 filler (non-[ti/di]) items. Non-[ti/di] syllables occurred with roughly equal frequency across novel blick test items. In the wug test, 48 novel singulars were created, with stem-final consonants appearing with equal frequency (8 for each consonant); familiar vowel-final stems were not included in this phase, and no novel vowel-final stems were included either. To make sure that there was no mismatching static information in stems in the wug test which allows us to use the same items for both language groups, none of the wug singulars contained [ti] or [di] stem-internally.

2.2. PARTICIPANTS. 69 native American English participants were recruited from the UCLA Psychology Pool, and each of them received course credit for their participation. There were 35 participants in the Across-the-board language, and 34 in the Derived-enivornment language. Another 16 participants were tested but were excluded for not having English as a first language (n = 1), having issues with playing the audio files (n = 4) and for not completing the task on the first attempt (n = 11; 4 in the Across-the-board language, and 7 in the Derived-environment language).

2.3. AUDIO AND VISUAL STIMULI. Audio stimuli were recorded by a female phonetically-trained native speaker of American English. She was instructed to ensure that voiced stops were voiced throughout the closure, and voiceless stops were always aspirated regardless of position in a word. She also ensured that stress was always placed on the initial syllable of the target word. Recordings were made in a sound-attenuated booth at a sampling rate of 22,050 Hz using PCQuirer (Scicon R & D 2015) using a Shure SM10A head-mounted microphone. Visual stimuli consisted of digital images that contained either a single object or animal or many objects or animals. Images were paired with the target words used in training, and in the wug phase.

2.4. PROCEDURE. Participants were tested online using Experigen (Becker & Levine 2014). In the training phase, listeners were presented first with a singular image that appeared on the left of the screen and heard a singular word for that object. Then they saw a plural image on the right, and heard the relevant plural word. Participants heard all the training words twice across two blocks, with presentation order randomized within blocks. This meant that Across-the-board learners had 104 trials (52 words X 2 blocks), whereas Derived-environment learners had 116 trials (58 words X 2 blocks) in total. Participants were told that they were going to be learning a new language and were encouraged to say the words out loud to help them learn the language.

Following the training phase, participants commenced the blick test where they performed phonotactic judgements on both familiar and novel singular items. This phase allowed us to probe what participants internalized about the static phonotactic patterns in the trained lexicon. Participants made a two-alternative forced choice on whether or not the word they heard was a possible word in the language they had just learned. Participants did not see images in this phase.

Familiar trials included filler words heard by both language groups, as well as [ti/di] words heard by the Derived-environment group. There were 36 novel trials, with 52 familiar ones resulting in 88 test trials in total.

After the blick phase, participants proceeded on to the wug test. Presentation of trials resembled the training phase. First, participants heard a singular word paired with a singular object, then they saw a plural object. Here, they heard two possible options for the plural: (one changing: [p/t]/[b/d] → [ʧ]/[ʤ] respectively, [ʧ]/[ʤ] → [t]/[d]; and one non-changing). Participants were instructed to choose the best word for the image that they were seeing. Order of plurals was counterbalanced such that participants heard changed plurals first on half the trials, and unchanged-plurals first on the other half. Participants did not hear any singular stems with stem-internal [ti] and [di]. There were 48 novel trials, with a total of 80 trials including 32 familiar filler words presented to both language groups in training.

2.5. RESULTS. We first discuss the results of the wug test. Here we are interested in the rate of choosing changed plurals depending on the stem-final consonant. Recall that on every trial in the wug test, listeners heard two options for the plural: a plural with an unchanged stem-final consonant, and one with a changed consonant. Whereas a changed plural for [p, b] and [t, d]-final stems contained [ʧ, ʤ] depending on voicing, a changed plural for a [ʧ, ʤ]-final stem contained [t, d]. Thus if participants successfully learned the alternation, we expect that they should only choose a changed plural for a [t, d]-final stems and not [p, b] or [ʧ, ʤ]-final stems.

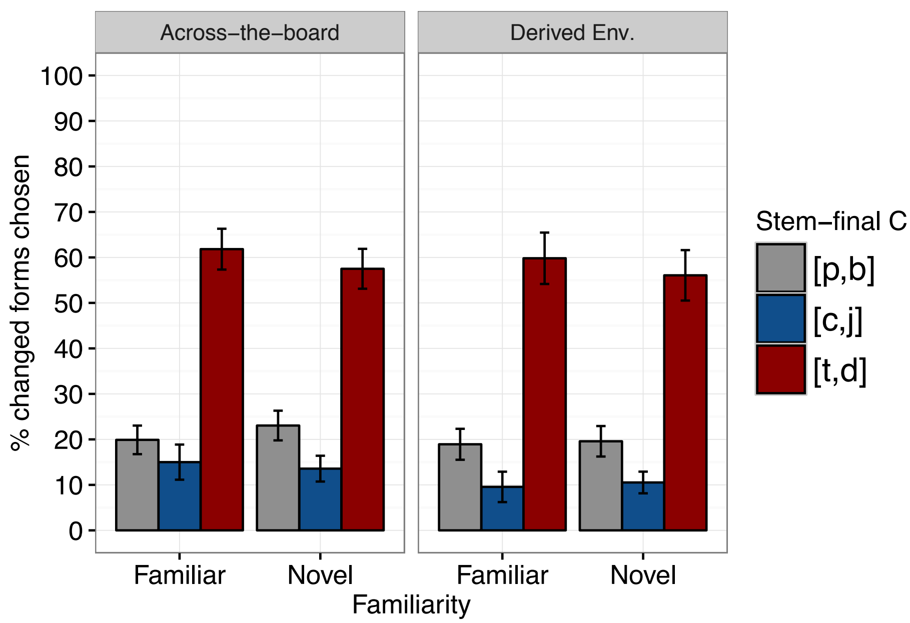

Figure 1. Experiment 1 wug test: Proportion of changed plurals selected by familiarity (familiar vs. novel) and stem-final consonant ([p,b],[ʧ,ʤ] ([c, j] in legend) & [t,d]). (L): Across-the-board language; (R): Derived-environment language

Proportion of changed plurals selected is shown in Figure 1. Overall, participants chose changed plurals more often with [t, d]-final stems than with other stem types, indicating that they successfully learned the alternation. Responses were modeled using mixed-effects logistic regression (see Jaeger 2008) in R (R Core Team 2015) using the *lme4* package (Bates et al. 2015). 1 additional participants data was excluded for only clicking "Yes" in the blick test. The model had Familiarity (reference = familiar), Language (reference = Across-the-board) and Word Type (reference = t/d), as well as two- and three-way interactions, as fixed factors. It also included random intercepts for subject and item, as well as random slopes for Word Type by subject and Language by item. This was the largest model to successfully converge. To determine the significance of a particular factor, likelihood ratio tests were conducted using the *anova()* function in R, by comparing two models in a subset relation.

The inclusion of the three-way interaction ($\chi^2(2) = 0.22$, $p = 0.90$) and the two-way interactions of Condition X Word Type ($\chi^2(2) = 3.22$, $p = 0.19$) and Language X Condition ($\chi^2(1) = 0.14$, $p = 0.71$) did not significantly improve the model. Crucially, the two-way interaction of Word Type X Language ($\chi^2(2) = 0.57$, $p = 0.75$) was not significant, indicating that the difference in the rate of selecting changed plurals based on stem-final consonant did not differ by Language. Additionally, the effect of Language was not significant ($\chi^2(2) = 0.41$, $p = 0.82$) and neither was the effect of Condition ($\chi^2(2) = 0.41$, $p = 0.82$). Importantly, the effect of Word Type significantly improved the model ($\chi^2(2) = 0.41$, $p = 0.82$). The results indicate that learners picked changed plurals significantly more often when the stem-final consonant was /t, d/ than when it was /p, b/ or /ʧ, ʤ/, as indicated by the negative model coefficients when compared to /ʧ, ʤ/ and /p, b/ stems as shown in Table 3. This means that participants successfully learned the alternation, although contrary to predictions, their success did not differ depending on the language which they were trained on.

Fixed Factors	Est.	*SE*	*z*-value	*p*	
Intercept (= t, d)	0.53	0.21	2.46	0.01	*
Word Type (= ʧ,ʤ)	-3.37	0.36	-9.40	< 0.001	***
Word Type (= p,b)	-2.44	0.30	-8.24	< 0.001	***

Table 3. Final model results for endorsement rate in wug test in Expt. 1.

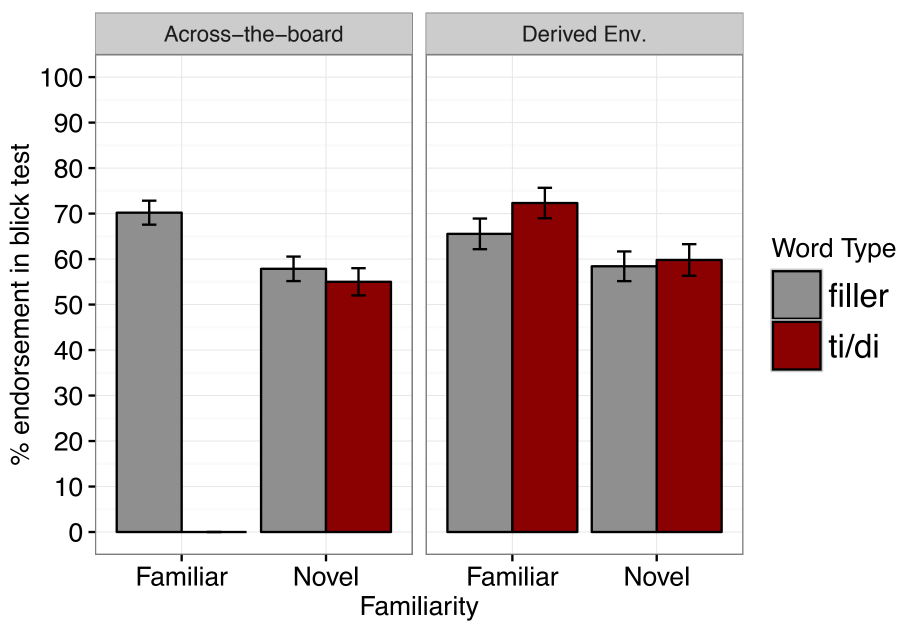

Figure 2. Experiment 1 blick test: Endorsement rate for singular words by familiarity (familiar vs. novel) and word type (filler vs. [ti/di]) (L): Across-the-board language; (R): Derived-environment language

Next, we examine the rate of endorsement for singular stems in the blick test to ascertain what static phonotacic generalizations participants in both language groups learned. Endorsement rates by familiarity (familiar vs. novel), word type (filler vs. [ti/di]) and language is shown in Figure 2. Notice that there were no familiar [ti/di] singulars for Across-the-board learners since they never heard forms with such sequences in training. Participants in the Across-the-board language did hear such forms at test (familiar [ti/di] stems for Derived-Environment learners), and for the purpose of the analysis all [ti/di] stems were coded as "novel" in the Across-the-board language. This meant that not only were there different items in novel [ti/di] items across both languages, but also that an entire combination of variables (familiar [ti/di] items) was missing in the Across-the-board language. This combination of factors made it difficult to analyze both languages together in a single analysis, and thus we analyzed these results separately by language. Here we are primarily interested in whether participants in each language differentiated between filler and [ti/di] singulars. Endorsement rate in the Across-the-board was modeled with Word Type as a fixed factor with three levels (Familiar-Filler, Novel-Filler and Novel-[ti/di]) with Novel-[ti/di] as the reference level, and random intercepts for subject and item, as well as random slopes for Word Type by subject. We expected that Across-the-board learners endorse fillers (familiar and novel) more often than [ti/di] words. That is, over and above any effect of unfamiliarity, we expected to still see higher endorsement rates for novel fillers compared to novel [ti/di] items. There was significant effect of Word Type as indicated by model comparison ($\chi^2(2)=14.8$, $p < 0.001$). Pair-wise comparisons were conducted using the *glht()* function of the *multcomp* package in R (Hothorn, Bretz & Westfall 2008) with the estimates of the model shown in Table 3. The statistical results show that while [ti/di] items were endorsed less often than familiar fillers, they were endorsed at the same rate as novel fillers. In contrast to the prediction, learners failed to infer a phonotactic constraint against singulars with [ti/di] sequences. That is, over and above the effect of unfamiliarity, there was no penalty for having [ti/di] sequences in stems since participants also endorsed novel fillers and [ti/di] stems at the a similar rate.

	Est.	*SE*	*z*-value	*p*	
[ti/di] vs. Familiar Filler	0.78	0.24	3.71	<0.001	*
[ti/di] vs. Novel Filler	0.14	0.24	0.57	0.84	
Familiar Filler vs. Novel Filler	-0.64	0.19	-3.36	0.002	**

Table 3. Final model results for endorsement rate in blick test for the Across-the-board language in Expt. 1.

For the Derived-environment language, endorsement rate was analyzed using a mixed-effects logistic regression with Familiarity (familiar vs. novel) and Word Type (filler vs. [ti/di]) as fixed effects and participant and item as random intercepts. Random slopes for Word Type and Language were also included in the model. This was the largest model to successfully converge. There was no significant two-way interaction of Familiarity by Word Type ($\chi^2(1) = 0.80$, $p = 0.37$), although there was a significant effect of Familiarity ($\chi^2(1) = 8.71$, $p = 0.003$) with novel items generally endorsed less often than familiar items regardless of Language and Word Type. Importantly, there was no significant effect of Word Type ($\chi^2(1) = 1.06$, $p = 0.30$). The results indicate then that participants endorsed filler and [ti/di] singulars at equal rates, regardless of familiarity, although they endorsed novel singulars less on the whole, similar to participants in the Across-the-board language.

2.6. DISCUSSION. Contrary to expectations, learners in both languages learned the alternation equally well. Interestingly, although learners in the Derived-environment language noticed [ti/di] singulars in training, they nonetheless learned the alternation as well as Across-the-board learners who did not hear such sequences at all. Despite this, Across-the-board learners failed to infer the phonotactic constraint against [ti] and [di], endorsing novel [ti/di] stems as often as novel filler stems. We also found a significant effect of familiarity across both languages. Given this, it is possible that our failure to find any difference in endorsement rate in the blick test was due to the fact that participants were responding in the blick test based entirely on whether or not they had heard the relevant forms in training. That is, it might be the case that they were just relying on pure memorization and recall in performing the task.

3. Experiment 2. In Experiment 2, participants were tested on the same languages except they were not tested on familiar singular forms in the blick test; they were only tested on novel unheard singulars. This was done in the hopes that participants would endorse forms not purely based on whether or not they had heard the exact word in training.

3.1. PARTICIPANTS. 43 native American English speakers were recruited from the UCLA Psychology Pool. All received course credit for their participation. Another 10 were tested but excluded due to having the wrong language background (n = 4), for issues with playing the sound files (n = 4) or for not completing the task on the first attempt (n = 2).

3.2. AUDIO AND VISUAL STIMULI. The same audio and visual stimuli were used in Experiment 2 as in Experiment 1.

3.3. PROCEDURE. The procedure in Experiment 2 was largely the same as in Experiment 1. In Experiment 2, however, participants were only tested on novel singular stems in the blick test, and not those heard in training. They were still tested on familiar and novel plurals in the wug test. In the blick test, participants were told that they had not heard any of the upcoming words previously, in the hopes that they would rely less of pure recall in the task.

3.4. RESULTS. As in Experiment 1, the proportion of changed plurals selected in the wug test was calculated by Familiarity, Word Type, and Language. This is shown in Figure 3. Another mixed-effects logistic regression was used to analyze the rate of change plural selection with the same fixed effects as above, with significance of particular factors determined by likelihood ration tests comparing models in a subset relation, starting with the fully-specified model.

The inclusion of the three-way ($\chi^2(2) = 0.41$, $p = 0.82$) and two-way interactions did not significantly improve the model (Word Type X Language: $\chi^2(2) = 2.27$, $p = 0.32$; Condition X Word Type: $\chi^2(2) = 4.22$, $p = 0.12$; Language X Condition: $\chi^2(2) = 0.05$, $p = 0.83$). Thus, the difference in the rate of selecting changed plurals by stem-final consonant did not differ by Language. The effect of Language was not significant ($\chi^2(2) = 0.41$, $p = 0.82$) and neither was the effect of Condition ($\chi^2(1) < 1$, $p = 0.98$). Importantly, the effect of Word Type significantly improved the model ($\chi^2(2) = 29.01$, $p < 0.001$). A summary of the final model is shown in Table 4. The results indicate that learners picked changed plurals significantly more often when the stem-final consonant was [t, d] than when it was [p, b] or [ʧ, ʤ]. This means that participants successfully learned the alternation, and importantly their success did not differ depending on the language which they were trained on. Thus, we successfully replicated the results of Experiment 1.

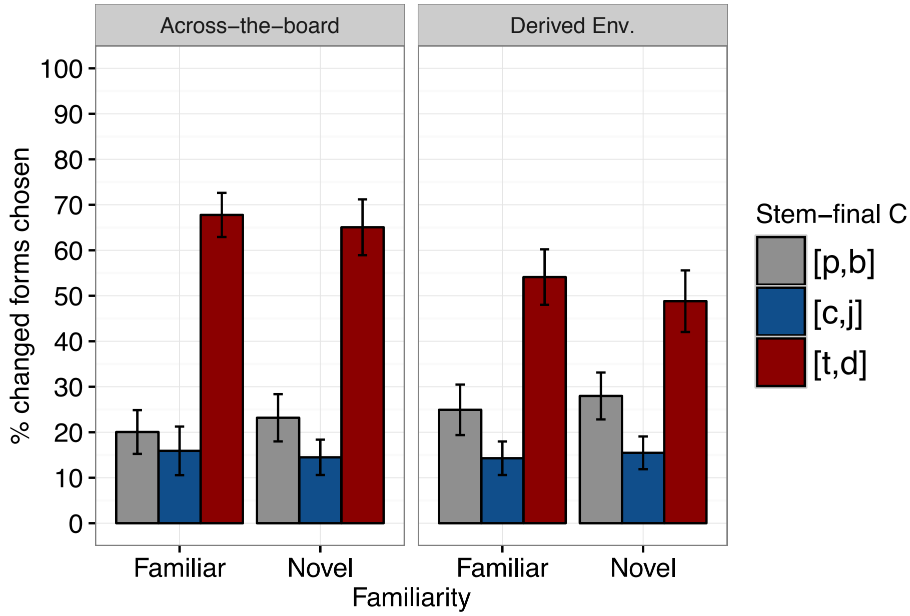

Figure 3. Experiment 2 wug test: Proportion of changed plurals selected by familiarity (familiar vs. novel) and stem-final consonant ([p,b],[ʧ,ʤ] ([c, j] in legend) & [t,d]). (L): Across-the-board language; (R): Derived-environment language

Fixed Factors	Est.	*SE*	*z*-value	*p*	
Intercept (= t, d)	0.71	0.29	2.47	0.01	*
Word Type (= ʧ,ʤ)	-3.36	0.56	-5.99	< 0.001	***
Word Type (= p,b)	-2.43	0.47	-5.19	< 0.001	***

Table 4. Final model results for endorsement rate in wug test in Expt. 2.

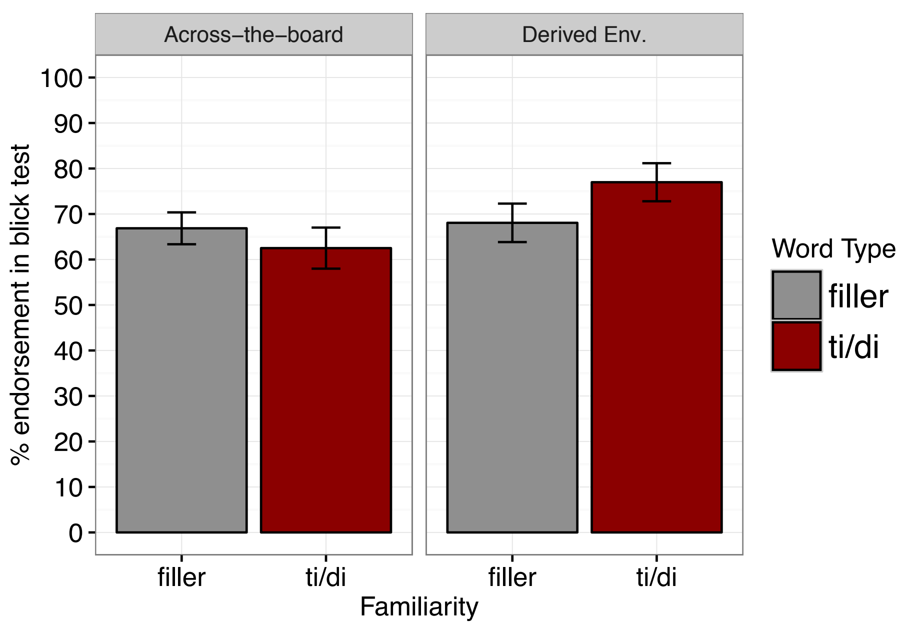

Figure 4. Experiment 2 blick test: Endorsement rate for singular words by word type (filler vs. [ti/di]) (L): Across-the-board language; (R): Derived-environment language.

Next we analyzed the endorsement rate of novel singulars in the blick test. Another mixed-effect logistic model was fit to the data with Word Type (reference = [ti/di]), Language (reference = Across-the-board) and their interaction as fixed effects, together with participant and singular word as random intercepts and random slopes for Word Type by participant and Language by singular word. There was a significant two-way interaction of Word Type X Language ($\chi^2(1) = 5.34$, $p = 0.02$). The final model estimates are shown in Table 5. The interaction is driven primarily by the fact that participants in the Derived-environment language endorsed [ti/di] items significantly more often than participants in the Across-the-board language ($\beta = 0.83$, $SE = 0.41$, $p = 0.04$). While Across-the-board learners failed to correctly differentiate between filler and [ti/di] singulars, they nonetheless endorsed [ti/di] singulars significantly less often than learners in the Derived-environment language. This suggests indirectly that learners' behavior differed depending on the language they were trained on.

Fixed Factors	Est.	*SE*	*z*-value	*p*	
(Intercept)	0.69	0.34	2.04	0.04	*
Word Type (= Filler)	0.24	0.32	0.74	0.46	
Language (= Derived env.)	0.83	0.41	2.04	0.04	*
Word Type X Language	-0.78	0.32	-2.44	0.01	*

Table 5. Final model results for endorsement rate in blick test in Expt. 2.

3.5. DISCUSSION. In Experiment 2, we replicated the general pattern of results we found in Experiment 1. Learners in both language groups learned the alternation equally well in the wug test. In the blick test, there was no difference between endorsement rates for filler and [ti/di] singulars within the Across-the-board language and the Derived-environment language. However, Across-the-board learners endorsed [ti/di] singulars less often than Derived-environment learners. This gives us some indirect evidence of Across-the-board learners' dispreference towards [ti/di] singulars when compared to Derived-environment learners.

4. General discussion & conclusion. Across two artificial grammar learning experiments, we examined the relationship between phonotactics and phonological alternations in learning. Specifically, we were interested in answering the question of whether or not alternation learning would be more difficult in a language with a derived-environment pattern, where these generalizations mismatch.

In Experiment 1, participants were trained either on a language with an Across-the-board pattern (matching static and dynamic generalizations) or one with a Derived-environment pattern (mismatching static and dynamic generalizations). In a departure from previous artificial grammar learning experiments, participants were tested on what they learned regarding both static phonotactic generalizations and phonological alternations. Surprisingly, participants in both languages learned the relevant alternation to the same extent. Participants, however, did not differ in their learning of phonotactics. Even in the Across-the-board language, learners did not infer a strong phonotactic constraint banning [ti] and [di] sequences, despite these sequences never appearing in the training data. We found the same pattern of results in Experiment 2 where participants were only asked to make well-formedness judgments on unseen novel singular forms. In Experiment 2, however, we found that Across-the-board learners endorsed [ti/di] singulars significantly less often than Derived-environment learners. This provides some indirect evidence of Across-the-board learners' dispreference for words with [ti] and [di] sequences when compared Derived-environment learners, although this dispreference was not robust given the lack of differentiation from filler words.

The results of our experiments are therefore unable to conclusively answer the question of whether or not a derived-environment pattern is more difficult to learn. Since participants in the Across-the-board language did not show robust learning of the relevant phonotactic constraint, we unfortunately lack a basis to compare the success with which participants learned the alternation in both languages. It may be the case that the lack of any difference in alternation learning we found here was entirely due to the fact that Across-the-board learners did not learn the crucial phonotactic constraint, and thus negating any facilitating effect matching phonotactics might have had.

The failure of successful learning in our experiment here stands in contrast to the successful learning of phonotactics in some previous studies (e.g. Warker & Dell 2006; Linzen & Gallagher 2014). The concern of these studies was solely on phonotactic learning, without phonological alternations. One possibility is that being trained on both singulars and plurals in a single training trial effectively highlights the alternation, and thus learners might fail to pay attention to any static phonotactic patterns. It is possible then that changing the presentation of the target words in training might better facilitate phonotactic learning. Another possibility might have to do with the type of phonotactic pattern involved. In both Warker and Dell (2006) and Linzen and Gallagher (2014), for example, participants were trained on phonotactic constraints that involved specific segments being licit in onsets (e.g. /f/ only occurs in onsets) and codas. These kinds of phonotactic patterns might be easier to learn in an artificial grammar setting compared to the kind of phonotactic pattern examined here. In the present case, learners had to infer a phonotactic gap without negative evidence, a task that might require much more training data than was presented here. Thus, it appears that different types phonotactic constraints are not learned at the same rate or with the same ease.

While we were not able to answer our main research question, a question we can answer is whether or not learning an alternation aids learners in noticing a phonotactic gap. The answer here seems to be no. Despite successfully learning the phonological alternation, Across-the-board learners nevertheless failed to extend that knowledge to static phonotactic judgments. Our results echo a recent finding by Pizzo (2015) which similarly failed to find a robust effect of alternation learning on phonotactic judgments. Pizzo's study examined whether or not learners extend a learned alternation (voicing or place assimilation) to static phonotactic judgments. While she found a significant extension of alternation learning onto phonotactic judgments in an explicit learning task with feedback, when using an implicit task without feedback, learners failed to show successful extension of this knowledge. Together, Pizzo's study and our current results suggest that there is no clear backwards effect of alternation learning on phonotactic judgments.

In sum, the results of our current study show that the learning of an alternation can proceed without supporting phonotactics. But our results raise further questions regarding the relationship between phonotactic and alternation learning, and relatedly how a grammar should encode these two kinds of phonological knowledge. Future work would seek to further examine this question in more detail, exploring some of the possibilities laid out above.

References

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. doi:http://dx.doi.org/10.18637/jss.v067.i01.

Becker, Michael, Nihan Ketrez & Andrew Nevins. 2011. The Surfeit of the Stimulus: Analytic Biases Filter Lexical Statistics in Turkish Laryngeal Alternations. *Language* 87(1). 84–125. doi:http://dx.doi.org/10.1353/lan.2011.0016.

Becker, Michael & Jonathan Levine. 2014. Experigen - an online experiment platform. http://becker.phonologist.org/experigen.

Berko, Jean. 1958. The Child's Learning of English Morphology. *WORD* 14(2-3). 150–177. doi:http://dx.doi.org/10.1080/00437956.1958.11659661.

Burzio, Luigi. 2011. Derived Environment Effects. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*. Blackwell. http://www.companiontophonology.com/subscriber/tocnode.html?id=g9781405184236_chunk_g978140518423690.

Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Cho, Taehong. 2001. Effects of Morpheme Boundaries on Intergestural Timing: Evidence from Korean. *Phonetica* 58(3). 129–162. doi:http://dx.doi.org/10.1159/000056196.

Cristià, Alejandrina, Jeff Mielke, Robert Daland & Sharon Peperkamp. 2013. Similarity in the generalization of implicitly learned sound patterns. *Laboratory Phonology* 4(2). doi:http://dx.doi.org/10.1515/lp-2013-0010.

Fountain, Amy V. 1998. An optimality theoretic account of Navajo prefixal syllables. Tucson: University of Arizona Ph.D. Dissertation.

Halle, Morris. 1978. Knowledge unlearned and untaught: What speakers know about the sounds of their language. *Linguistic theory and psychological reality*, 294–303. Cambridge, MA: MIT Press.

Hayes, Bruce. 2004. Phonological acquisition in Optimality Theory: the early stages. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in phonological acquisition*, 158–203. Cambridge University Press.

Hayes, Bruce, Péter Siptár, Kie Zuraw & Zsuzsa Londe. 2009. Natural and Unnatural Constraints in Hungarian Vowel Harmony. *Language* 85(4). 822–863. doi:http://dx.doi.org/10.1353/lan.0.0169.

Hayes, Bruce & James White. 2013. Phonological Naturalness and Phonotactic Learning. *Linguistic Inquiry* 44(1). 45–75. doi:http://dx.doi.org/10.1162/LING_a_00119.

Hayes, Bruce & Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39(3). 379–440. doi:http://dx.doi.org/10.1162/ling.2008.39.3.379.

Hothorn, Torsten, Frank Bretz & Peter Westfall. 2008. Simultaneous Inference in General Parametric Models. *Biometrical Journal* 50(3). 346–363. doi:http://dx.doi.org/10.1002/bimj.200810425.

Iverson, Gregory K. & Deirdre W. Wheeler. 1988. Blocking and the Elsewhere condition. In Michael Hammond & Michael Noonan (eds.), *Theoretical morphology: Approaches in modern linguistics*, 325–338. San Diego: Academic Press.

Jaeger, T. Florian. 2008. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of memory and language* 59(4). 434–446. doi:http://dx.doi.org/10.1016/j.jml.2007.11.007.

Jarosz, Gaja. 2006. Rich lexicons and restrictive grammars - Maximum likelihood learning in Optimality Theory. Johns Hopkins University Ph.D. Dissertation.

Kari, James M. 1976. *Navajo verb prefix phonology*. New York: Garland Press.

Kenstowicz, Michael & Charles Kisseberth. 1977. *Topics in phonological theory*. Academic Press.

Kiparsky, Paul. 1973. Phonological representations. In Osamu Fujimura (ed.), *Three dimensions of linguistic theory*, 1–135. Tokyo: TEC.

Kiparsky, Paul. 1993. Blocking in nonderived environments. In Ellen Kaisse & Sharon Hargus (eds.), *Phonetics and Phonology 4: Studies in lexical phonology*, 277–313. San Diego: Academic Press.

Linzen, Tal & Gillian Gallagher. 2014. The timecourse of generalization in phonotactic learning. In John Kingston, Claire Moore-Cantwell, Joe Pater & Robert Staub (eds.), *Proceedings of Phonology 2013*. Washington, DC: Linguistic Society of America.

Martin, Andrew. 2011. Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language* 87(4). 751–770. doi:http://dx.doi.org/10.1353/lan.2011.0096.

McDonough, Joyce. 1991. On the representation of consonant harmony in Navajo. *Proceedings of the West Coast Conference on Formal Linguistics (WCCFL)*, vol. 10, 319–335.

McDonough, Joyce. 2003. *The Navajo sound system*. Dordrecht: Kluwer.

Paster, Mary. 2013. Rethinking the "duplication problem." *Lingua* 126. 78–91. doi:http://dx.doi.org/10.1016/j.lingua.2012.11.015.

Pater, Joe & Anne-Michelle Tessier. 2005. Phonotactics and alternations: Testing the connection with artificial language learning. In Kathryn Flack & Shigeto Kawahara (eds.), *University of Massachusetts Occasional Papers in Linguisitcs [UMOP]*, vol. 31, 1–16.

Pizzo, Presley. 2015. Investigating properties of phonotactic knowledge through web-based experimentation. University of Massachusetts Amherst Ph.D. Dissertation.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Cambridge, MA: Blackwell Publishers.

Prince, Alan & Bruce Tesar. 2004. Learning phonotactic distributions. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in phonological acquisition*, 245–291. Cambridge University Press.

R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sapir, Edward & Harry Hoijer. 1967. *The phonology and morphology of the Navaho language*. Berkeley: University of California Press.

Scicon R & D. 2015. *PCQuirer*. Los Angeles, CA.

Skoruppa, Katrin & Sharon Peperkamp. 2011. Adaptation to Novel Accents: Feature-Based Learning of Context-Sensitive Phonological Regularities. *Cognitive Science* 35(2). 348–366. doi:http://dx.doi.org/10.1111/j.1551-6709.2010.01152.x.

Tesar, Bruce & Alan Prince. 2007. Using phonotactics to learn phonological alternations. *Proceedings of the Thirty-Ninth Conference of the Chicago Linguistics Society, vol. II: The panels*, 209–237.

Warker, Jill A. & Gary S. Dell. 2006. Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology. Learning, Memory, and Cognition* 32(2). 387–398. doi:http://dx.doi.org/10.1037/0278-7393.32.2.387.

White, James. 2014. Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130(1). 96–115. doi:http://dx.doi.org/10.1016/j.cognition.2013.09.008.

Wilson, Colin. 2006. Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* 30(5). 945–982. doi:http://dx.doi.org/10.1207/s15516709cog0000_89.

Wolf, Matthew. 2008. Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model. University of Massachusetts, Amherst Ph.D. Dissertation.