Substantive bias in phonotactic learning: Positional extension of an obstruent voicing contrast

Eleanor Glewwe
University of California, Los Angeles

1 Introduction

A current subject of debate in phonological theory is the extent to which the phonological typology is shaped by synchronic learning biases. Synchronic bias, also called analytic bias (Moreton 2008), refers to the notion that learners are biased toward acquiring certain phonological systems (the attested ones) over others (the unattested ones). Two types of synchronic bias have been discussed and tested in the literature: complexity bias, a bias against formally complex patterns, and substantive bias (also called naturalness bias), a bias against phonetically unnatural patterns. This paper investigates whether phonetic naturalness biases phonotactic learning. I approach this question by testing whether learners reproduce an attested and phonetically-motivated phonotactic implicational in an artificial grammar learning (AGL) experiment. The results do not fit the pattern predicted by the phonotactic implicational and thus provide no evidence for substantive bias, but the experiment does yield partial support for complexity bias.

2 Background

2.1 Past research on synchronic learning biases

A number of studies have uncovered evidence for complexity bias by finding that learners acquire featurally simpler patterns better than patterns that are featurally more complex (Moreton 2008, Hayes et al. 2009, Skoruppa & Peperkamp 2011, Moreton 2012).

Other studies have found evidence for substantive bias in that learners prefer to acquire phonetically natural patterns and underlearn phonetically unnatural patterns (Wilson 2006, Becker, Ketrez & Nevins 2011, Becker, Nevins & Levine 2012, Finley 2012, Hayes & White 2013, White 2013). However, these studies do not all yield unambiguous support for substantive bias. While Becker, Ketrez & Nevins (2011) and Hayes & White (2013) claim to have found effects of naturalness bias, their results can be reinterpreted as showing effects of complexity bias instead. Additionally, Wilson's (2006) pattern of results is not fully consistent with a naturalness bias account since in one experimental condition learners extended a palatalization process in both a natural direction and an unnatural direction.

Most studies probing synchronic bias in phonological learning have used an artificial grammar learning (AGL) paradigm. Moreton & Pater's (2012a, 2012b) review of work in this area found numerous results supporting complexity bias but relatively few supporting naturalness bias. They therefore concluded that there is fairly robust evidence for complexity bias but scant evidence for substantive bias.

2.2 Synchronic biases in phonotactic learning

Investigations of substantive bias have focused mostly on alternations. Only a few studies have tested for substantive bias in phonotactic learning. Skoruppa & Peperkamp (2011) taught subjects artificial variants of French that featured vowel harmony (Harmonic French) or vowel disharmony (Disharmonic French). Both the harmony and disharmony patterns were static generalizations; there were no alternations. While vowel harmony is typologically common and phonetically motivated and vowel disharmony is less common and not obviously phonetically motivated, subjects learned Harmonic French and Disharmonic French equally well, as evidenced by their ability to identify words that could belong to the language they had been exposed to. The study thus yielded no evidence for substantive bias. Skoruppa & Peperkamp did find a complexity bias: subjects performed worse on a variant of French featuring a mixed pattern of harmony and disharmony whose expression in features was more complex than the featural expression of the generalizations in Harmonic and Disharmonic French.

Hayes & White (2013) used the UCLA Phonotactic Learner (Hayes & Wilson 2008) to uncover natural and unnatural phonotactic constraints on the English lexicon. The constraints had similar weights (i.e. they were of similar strength in the lexicon), but the unnatural ones had little typological or phonetic support. Hayes & White found that English speakers underlearn the unnatural phonotactic constraints, as evidenced by a wug test. That is, English speakers had not internalized unnatural phonotactic constraints that were nevertheless supported by the lexicon, suggesting an effect of substantive bias. It is possible that the effect that emerged in this study reflected complexity bias, however, because the unnatural constraints may have been more complex than the natural ones. The natural constraints related two instances of the same feature more often than the unnatural constraints did. Additionally, the unnatural constraints were more likely to concern dependencies between vowels and consonants. As a result, the unnatural constraints may have been underlearned due to their complexity rather than due to their unnaturalness.

Myers & Padgett (2014) conducted a series of experiments examining generalization of final devoicing from the domain of the phrase to the domain of the word. In Experiment 1, they taught subjects either a natural phonotactic restriction against phrase-final voiced obstruents or an unnatural phonotactic restriction against phrase-final voiceless obstruents. Both phonotactic systems were learned to the same degree, so the experiment provided no support for substantive bias.

Finally, Greenwood (2016) carried out an experiment that compared the learnability of a language with a natural phonotactic restriction against word-final voiced obstruents and a language with an unnatural phonotactic restriction against word-final voiceless obstruents. In the casual speech condition, in which the stimuli were produced naturally, both languages were learned to the same degree, meaning that there was no effect of substantive bias. In the careful speech condition, in which the stimuli were hyperarticulated, the unnatural language that disallowed word-final voiceless obstruents was learned *better* than the natural language that disallowed word-final voiced obstruents. This pattern is the opposite of the one that would have constituted evidence for substantive bias. None of Greenwood's results support the idea that natural phonotactic generalizations are learned better than unnatural ones.

Summing up, Skoruppa & Peperkamp (2011), Myers & Padgett (2014), and Greenwood (2016) all failed to find effects of substantive bias on phonotactic learning in AGL experiments. Based on evidence from an English wug test, Hayes & White (2013) argue that naturalness does bias phonotactic learning, but their results may in fact support complexity bias. Skoruppa & Peperkamp also found an effect of complexity bias on phonotactic learning.

I take a different approach to testing for substantive bias in phonotactic learning by investigating not just phonotactic restrictions but a phonotactic implicational about the existence of contrasts in different positions. Given evidence for a contrast in one position, what do learners assume about that contrast's existence in other positions? I hypothesized that implicitly asking learners to compare the existence of contrasts across positions might cause an effect of substantive bias to emerge when simply testing the learnability of a specific phonotactic constraint does not.

2.3 The phonotactic implicational

The phonotactic implicational I tested is the following: If a language contrasts voicing in obstruents word-finally (e.g. /ap/ vs. /ab/), it will contrast voicing in obstruents word-initially (e.g. /pa/ vs. /ba/), but not necessarily vice versa. This implicational has a phonetic motivation rooted in perception. Steriade (1997) argued that cues to obstruent voicing are more abundant word-initially than word-finally; in particular, VOT is an available cue word-initially but not word-finally. As a result, voiced and voiceless obstruents should be more perceptually similar (that is, harder to distinguish) at the end of a word than at the beginning of a word. Thus if languages are phonetically natural, a language that has an obstruent voicing contrast word-finally, where it is harder to perceive, should also have that voicing contrast word-initially, where it is easier to perceive. On the other hand, there is no expectation from phonetic naturalness that a language with an obstruent voicing contrast word-initially should also have that voicing contrast word-finally. In addition to being phonetically motivated, the implicational is supported by the phonological typology; languages do conform to it (Steriade 1997). I conducted an experiment to test whether this phonotactic implicational is reproduced in artificial grammar learning.

3 Experiment: Positional extension of a voicing contrast

3.1 Method

The phonotactic learning experiment exposed subjects to an obstruent voicing contrast in either word-initial or word-final position and tested whether they extended the contrast to the other position. If participants' behavior is consistent with the phonotactic implicational about voicing contrasts, there should be asymmetrical extension: subjects exposed to the contrast word-finally, where it is less perceptible, should extend the contrast to word-initial position, where it is more perceptible, but not so much the other way around. This result would constitute evidence for a phonetic naturalness (i.e. substantive) bias in phonotactic learning.

3.1.1 Conditions

There were four training conditions defined on two dimensions: *Trained Contrast Position*, that is, whether the language exhibited an obstruent voicing contrast in word-initial or word-final position, and *Trained Neutralization Value*, that is, whether the language exhibited voiced or voiceless obstruents in the position without a contrast (or put another way, whether obstruents "neutralized" to voiced or voiceless in the position without a contrast). The properties of the four training conditions are shown in Table 1.

Table 1: The four training conditions.

Trained Contrast Position	Trained Neutralization Value	Condition	#T	#D	T#	D#
Final	**Voiced**	**#D…{T, D}# (*#T)**	✗	✓	✓	✓
	Voiceless	***#T…{T, D}# (*#D)***	✓	✗	✓	✓
Initial	Voiced	#{T, D}…D# (*T#)	✓	✓	✗	✓
	Voiceless	*#{T, D}…T# (*D#)*	✓	✓	✓	✗

The condition names, given in Table 1, express what types of obstruents occurred in which positions in each condition. For instance, in the **#D…{T, D}#** condition, only voiced stops occurred word-initially while both voiceless and voiced stops occurred word-finally. In the final contrast conditions (bolded), subjects heard both voiced and voiceless stops word-finally but only voiced or voiceless stops word-initially. The initial contrast conditions (not bolded) are the mirror image: subjects heard both voiced and voiceless stops word-initially but only voiced or voiceless stops word-finally. The conditions in which stops neutralized to voiceless in the position without the contrast are italicized, and the conditions in which stops neutralized to voiced are non-italicized. One can also define the four conditions by what obstruent voicing and position combination is excluded. For example, in the **#D…{T, D}#** condition there are no initial voiceless stops (i.e. *#T). These condition-defining phonotactic constraints are given in parentheses after the condition names in Table 1.

3.1.2 Materials

The training and test items were all nonce words of the shape $C_1VC_2VC_3$. Either C_1 or C_3 was a stop drawn from [p t k b d g]. The other two Cs were sonorants drawn from [m n l r j w] (the glides [j] and [w] did not occur word-finally, and no item contained a [ji] or [wu] sequence). In each word, all three Cs were different. The vowels were drawn from [i ɑ u]. Bilabial, alveolar, and velar stops were equally represented across positions and voicing categories in each condition and in both the training and test phases. In both training and test, half of the items were members of a minimal pair (in the training phase, this was only the case for items with stops in the position with an obstruent voicing contrast). This was to encourage participants to notice that voicing was contrastive. Additionally, in both training and test, half of the items were iambs and half were trochees. Stress did not correlate with the position that featured the obstruent voicing contrast or with the syllable in which the stop (always either word-initial or word-final) occurred. This was to prevent subjects from associating the obstruent voicing contrast with stress instead of position and to prevent stress from drawing undue attention to the stops.

Sample training items in the *#{T, D}…T#* condition are shown in Table 2. In this condition, participants were trained on words with voiced and voiceless stops in initial position but only voiceless stops in final position. This training language can be thought of as a final devoicing language, though as this is a phonotactic learning experiment subjects were not exposed to any alternations.

Table 2: Sample training items in the *#{T, D}…T# (*D#)* condition.

#T	#D	T#	D#
pímir **t**ilár **k**awám …	**b**ímir **d**irín **g**awám …	míwi**p** laní**t** nuwá**k** …	

The stimuli were recorded by a phonetically-trained female native speaker of American English. Voiced stops were fully voiced, voiceless stops were aspirated, and word-final stops were released. The stimuli were produced in isolation as if each item was a sentence unto itself. The stimuli were recorded in a soundproof booth using a Shure SM-10A head-mounted microphone plugged into an

XAudioBox. The recordings were done in the program PCQuirerX with a sampling rate of 22,050 Hz.

3.1.3 Procedure

The experiment was conducted online using Experigen (Becker & Levine 2013). Participants were instructed to wear headphones and to do the experiment in a quiet room. They saw written instructions informing them that they would be listening to some words of a new language. They would then be presented with additional words and asked whether those words sounded like they could also be from the language they heard in the first part of the experiment. Before the experiment proper began, they were prompted to play two sound files and type in the English word they had heard. The first word, *pad*, could be played multiple times. The second word, *bat*, could only be played once. These test words were included to try to ensure participants were doing the experiment under acceptable listening conditions.

The experiment began with the training phase. Participants proceeded at their own pace through two blocks of the same 36 training items. The order of the training items was randomized within each block for each participant. In each training trial, participants saw an image and clicked a button to hear the word for that image. Each sound file could only be played once. Participants then clicked to continue to the next training trial. They were encouraged to say the words out loud to help them learn the language. After the training phase, written instructions appeared telling participants that they would hear some additional words and should make their best guess as to whether each word sounded like it could also be a word from the language they had just listened to. There were no images in the test phase. In each test trial, participants clicked a button to hear the sound file. They then had to click Yes or No to indicate whether they thought the word sounded like it could be from the language they had just listened to. There was a single test block consisting of 48 test items. Their order was randomized for each participant. At the end of the experiment, participants completed a demographic survey.

Table 3: Sample test items for each training condition.

Condition	Familiar Conforming	Novel Conforming	Novel Nonconforming
#D…{T, D}# (*#T)	nimá**p**	rínu**p**	**p**írum
#T…{T, D}# (*#D)	nimá**p**	rínu**p**	**b**írum
#{T, D}…D# (*T#)	**k**awám	**k**ámir	múli**k**
*#{T, D}…T# (*D#)*	**k**awám	**k**ámir	múli**g**

There were three types of items in the test phase. *Familiar Conforming* items contained stops whose voicing and position conformed to the trained pattern (e.g. #T, #D, and T# in the *#{T, D}…T#* condition) and were words that were heard in training. *Novel Conforming* items also contained stops whose voicing and position conformed to the trained pattern, but these words had not been heard in training. Finally, *Novel Nonconforming* items featured the voicing and position combination not heard in training (e.g. D# in the *#{T, D}…T#* condition). Table 3 gives sample test items for each of the four training conditions.

3.1.4 Participants

The participants were native English speakers recruited through the UCLA Psychology Subject Pool. I excluded participants who were not native English speakers, had taken more than one linguistics class, reported a history of speech or hearing impairments, gave an incorrect response to either of the two test words that preceded the experiment, or accepted all test items. After exclusions, there were 149 participants divided among the four conditions. The number of participants per condition ranged from 33 to 41.

3.2 Predictions

Different hypotheses make distinct predictions about the results of the experiment. Participants' acceptance rates of Novel Nonconforming items, relative to Novel Conforming items, indicate whether they have extended the obstruent voicing contrast to a new position in a given condition. For instance, if subjects in the *#{T, D}…T#* condition accept test items with word-final voiced stops (D#), they have extended the word-initial obstruent voicing contrast they encountered in training to word-final position.

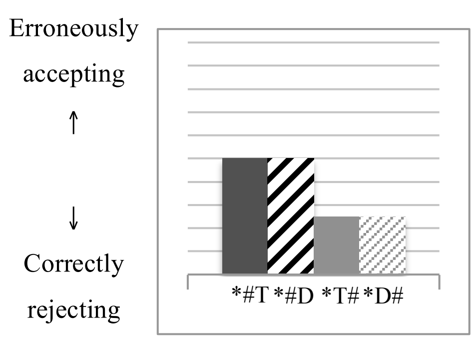

Figure 1: Predicted relative acceptance rates of Novel Nonconforming items by condition according to the position-based substantive bias hypothesis.

Consider first the position-based substantive bias hypothesis. Recall the phonotactic implicational: an obstruent voicing contrast in final position entails an obstruent voicing contrast in initial position, but not vice versa. Behavior consistent with the implicational would be asymmetrical extension: subjects exposed to the voicing contrast word-finally should extend it to word-initial position more than subjects exposed to the voicing contrast word-initially extend it to word-final position. Extension of the voicing contrast to the other position manifests as erroneously accepting Novel Nonconforming items, so subjects trained on the contrast word-finally should more readily accept their Novel Nonconforming items (#T and #D, depending on the condition) than subjects trained on the contrast word-initially accept their Novel Nonconforming items (T# and D#, depending on the condition). This pattern of results is illustrated in the schematic graph in Figure 1, which shows the predicted relative acceptance rates of Novel Nonconforming items among the four conditions according to the position-based substantive bias hypothesis. Each condition is identified by the type of item that is Novel Nonconforming (i.e. not attested in training) in that condition.

There is a second, voicing-based substantive bias hypothesis. Since voiced obstruents are more marked than voiceless obstruents (Greenberg 1978), this hypothesis predicts more extension from voiced obstruents to voiceless obstruents than from voiceless obstruents to voiced obstruents. That is, subjects trained on a neutralizing-to-voiced language should more readily accept their Novel Nonconforming items, which are voiceless (#T or T#), than subjects trained on a neutralizing-to-voiceless language accept their Novel Nonconforming items, which are voiced (#D or D#). In other words, there should be higher acceptance rates of Novel Nonconforming items in the conditions where Novel Nonconforming items feature a voiceless stop. This pattern of results is illustrated in the schematic graph in Figure 2.

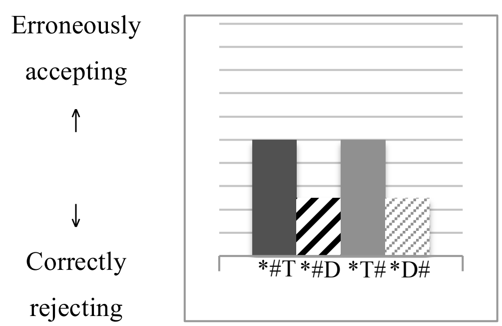

Figure 2: Predicted relative acceptance rates of Novel Nonconforming items by condition according to the voicing-based substantive bias hypothesis.

If both position-based and voicing-based substantive biases are at work, we predict more extension from word-final position to word-initial position as well as more extension from voiced obstruents to voiceless obstruents.

Finally, there is a potential prediction from complexity bias. Due to the presence of sonorant consonants in the training items, the phonotactic constraint that is needed to exclude Novel Nonconforming items in the neutralizing-to-voiceless conditions could be more complex than the constraint needed in the neutralizing-to-voiced conditions. To see how this works, we can compare two conditions that differ only in Trained Neutralization Value. In the #{T, D}…D# training condition, a neutralizing-to-voiced condition, subjects hear words that end in voiced sonorants (e.g. *kawá**m***) and words that end in voiced obstruents (e.g. *míwi**b***) but no words that end in voiceless obstruents (e.g. *míwi**p***). Thus to exclude the type of item they did not hear, subjects could posit the single-feature constraint *[-voice]#. The neutralizing-to-voiceless counterpart of the #{T, D}…D# condition is the #{T, D}…T# condition. In this condition, subjects hear words that end in voiced sonorants (e.g. *kawá**m***) and words that end in voiceless obstruents (e.g. *míwi**p***) but no words that end in voiced obstruents (e.g. *míwi**b***). To exclude words ending in voiced obstruents while not also excluding words ending in voiced sonorants, subjects in this condition must posit the two-feature constraint *[-son, +voice]#. The same is true for the other pair of conditions: subjects may learn a one-feature constraint to exclude voiceless obstruents in the neutralizing-to-voiced condition while they must learn a two-feature constraint to exclude voiced obstruents in the corresponding neutralizing-to-voiceless condition. Because it is potentially harder to learn to exclude voiced obstruents, the complexity bias hypothesis predicts more extension from voiceless obstruents to voiced obstruents than from voiced obstruents to voiceless obstruents. Note that this prediction is precisely the opposite of the voicing-based substantive bias prediction described above. The pattern of results predicted by complexity bias are illustrated in the schematic graph in Figure 3: Novel Nonconforming items with voiceless stops should be accepted less often (i.e. correctly rejected more often) than Novel Nonconforming items with voiced stops.

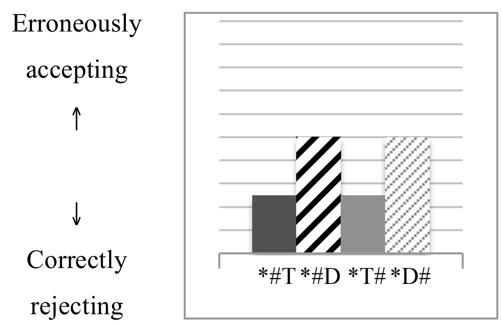

Figure 3: Predicted relative acceptance rates of Novel Nonconforming items by condition according to the complexity bias hypothesis.

If no synchronic biases are at work, we expect similar acceptance rates of Novel Nonconforming items across all four conditions.

3.3 Results

Figure 4 shows the acceptance rates of the three types of test items across conditions. A mixed-effects logistic regression fit to the conforming items (Familiar Conforming and Novel Conforming) with response (accept or reject) as the dependent variable; Familiarity (familiar vs. novel), Trained Contrast Position, Trained Neutralization Value, and their 2- and 3-way interactions as fixed effects; and random intercepts for subject and item yielded a significant main effect of

Familiarity: Familiar Conforming items were accepted more often than Novel Conforming items (β = 1.177; p < 0.001). This is expected: subjects should be better at accepting conforming items they heard in the training phase than conforming items that are new.

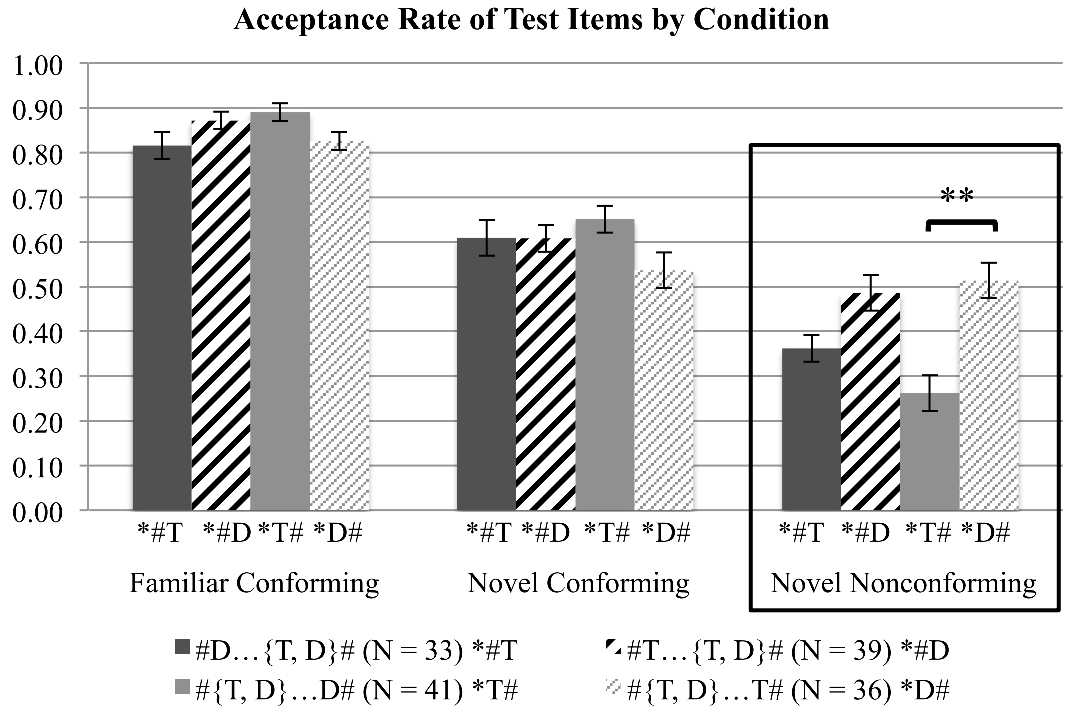

Figure 4: Acceptance rates of test items by condition. Acceptance rates of Novel Nonconforming items are framed and may be informally compared to the schematic graphs in Figures 1–3.

The acceptance rates of Novel Conforming items were above chance in all four conditions. This was determined by fitting mixed-effects logistic regressions to the novel items (Novel Conforming and Novel Nonconforming) and changing the reference levels of the factors Trained Contrast Position and Trained Neutralization Value so that in each of the four models the intercept represented the acceptance rate for Novel Conforming items in one of the four conditions. The intercept was significantly above chance in all four models, showing that subjects accepted Novel Conforming items at above chance levels in all four conditions. This means that in all conditions subjects were able to generalize to novel items that fit their training language, showing that they had learned the pattern they had been trained on. To check that acceptance rates of Novel Conforming items did not differ significantly among the four conditions, I fit a mixed-effects logistic regression just to Novel Conforming items with Condition as the fixed effect. There was no significant main effect of Condition, and post-hoc pairwise comparisons confirmed that no two conditions had significantly different acceptance rates of Novel Conforming items. Consequently, we can compare the acceptance rates of Novel Nonconforming items to test for the different biases reviewed in Section 3.2.

A mixed-effects logistic regression was fit to the Novel Nonconforming items with response (accept or reject) as the dependent variable, Condition as the fixed effect, and random intercepts for subject and item. I then conducted post-hoc pairwise comparisons (Tukey method) of the acceptance rates of Novel Nonconforming items to determine which pairs of conditions had significantly different acceptance rates.

Recall that the position-based substantive bias hypothesis predicts that subjects trained on the obstruent voicing contrast word-finally should accept Novel Nonconforming items more (i.e. extend the contrast to initial position more) than subjects trained on the voicing contrast word-initially do. This translates to greater acceptance of Novel Nonconforming items in the **#D…{T, D}#** condition than in the #{T, D}…D# condition and greater acceptance of Novel Nonconforming items in the ***#T…{T, D}#*** condition than in the *#{T, D}…T#* condition. Table 4 shows the two differences predicted by substantive bias as well as the actual direction of the differences and the *p*-values for those actual differences.

Table 4: Pairwise differences in acceptance rates of Novel Nonconforming items that test the position-based substantive bias hypothesis.

Predicted Difference	**Actual Difference**	***p***
#D…{T, D}# > #{T, D}…D#	**#D…{T, D}#** > #{T, D}…D#	0.528
#T…{T, D}# > *#{T, D}…T#*	***#T…{T, D}#*** < *#{T, D}…T#*	0.975

As Table 4 shows, the acceptance rate of Novel Nonconforming items was higher in the **#D…{T, D}#** condition than in the #{T, D}…D# condition, but not significantly so. The difference in acceptance rates of Novel Nonconforming items between the ***#T…{T, D}#*** condition and the *#{T, D}…T#* condition was actually in the direction opposite the predicted one, but in any case it was not significant. Thus neither of the position-based substantive bias hypothesis predictions were borne out.

Complexity bias predicts that subjects trained on a neutralizing-to-voiceless language will have a harder time learning to reject Novel Nonconforming items (i.e. voiced stops) than subjects trained on a neutralizing-to-voiced language. That is, the neutralizing-to-voiceless conditions should exhibit higher acceptance rates of Novel Nonconforming items than the corresponding neutralizing-to-voiced conditions. This translates to greater acceptance of Novel Nonconforming items in the ***#T…{T, D}#*** condition than in the **#D…{T, D}#** condition and greater acceptance of Novel Nonconforming items in the *#{T, D}…T#* condition than in the #{T, D}…D# condition. (Recall that the voicing-based substantive bias hypothesis predicts precisely the opposite differences.) Table 5 shows the two differences predicted by complexity bias, the actual direction of the differences, and the *p*-values of the pairwise comparisons that test those differences.

Table 5: Pairwise differences in acceptance rates of Novel Nonconforming items that test the complexity bias hypothesis.

Predicted Difference	Actual Difference	*p*
#T…{T, D}# > **#D…{T, D}#**	***#T…{T, D}#*** > **#D…{T, D}#**	0.408
#{T, D}…T# > #{T, D}…D#	*#{T, D}…T#* > #{T, D}…D#	0.004 **

As shown in Table 5, within both the initial contrast conditions and the final contrast conditions, Novel Nonconforming items were accepted more in the neutralizing-to-voiceless condition than in the neutralizing-to-voiced condition. Thus the predicted differences were in the expected direction in both cases. However, only for the initial contrast conditions was the difference significant. The acceptance rate of Novel Nonconforming items in the *#{T, D}…T#* condition was significantly higher than in the #{T, D}…D# condition. While the acceptance rate of Novel Nonconforming items in the ***#T…{T, D}#*** condition was higher than in the **#D…{T, D}#** condition, the difference was not significant. In other words, one of the complexity bias predictions was borne out, but the other was not.

4. Discussion

The phonotactic implicational I set out to test (i.e. an obstruent voicing contrast word-finally entails an obstruent voicing contrast word-initially, but not vice versa) was not reproduced in the present experiment. There was not greater extension of the obstruent voicing contrast from word-final position to word-initial position than from word-initial position to word-final position. The substantive bias hypothesis was therefore not supported.

Instead, subjects trained on a voicing contrast that neutralized to voiceless stops in one position seemed to extend to voiced obstruents in that position more than subjects trained on a voicing contrast that neutralized to voiced stops in one position extended to voiceless obstruents in that position. This is the opposite of the behavior we expect based on the relative markedness of voiced and voiceless stops. Given, the presence of (voiced) sonorants in the training items, though, this result can be explained by a complexity bias. In the neutralizing-to-voiced conditions, **#D…{T, D}#** and #{T, D}…D#, participants could respectively learn the one-feature constraints *#[-voice] and *[-voice]# in order to reject Novel Nonconforming items. In contrast, in the neutralizing-to-voiceless conditions, ***#T…{T, D}#*** and *#{T, D}…T#*, participants had to learn the more complex two-feature constraints *#[-son, + voice] and *[-son, + voice]#, respectively, to reject Novel Nonconforming items. While this complexity bias account does provide an explanation for my experimental results, it depends on English sonorants having an active [+voice] feature, which is a matter of debate (see for instance Steriade 1987).

That said, this surprising result whereby the phonotactic equivalent of a final voicing language is learned better than the phonotactic equivalent of a final devoicing language is not unprecedented. The unexpected results of Greenwood's (2016) phonotactic learning experiment exhibit a similar pattern. Greenwood's experiment had two conditions: in one condition, participants were exposed to a language with a phonotactic restriction against word-final voiceless obstruents, and in the other participants were exposed to a language with a phonotactic restriction against word-final voiced obstruents. In Greenwood's experiments, obstruents were fricatives and affricates rather than stops, so her experiments' phonotactic constraints can be summarized as *S# and *Z#. Crucially, as in my experiment, the training items in both of Greenwood's conditions also included words that ended in (voiced) sonorants. Thus in her experiment participants in the *S# condition could have posited the one-feature constraint *[-voice]# to correctly exclude words ending in voiceless obstruents while participants in the *Z# condition would have had to posit the two-feature constraint *[-son, +voice]# to correctly exclude words ending in voiced obstruents but not words ending in voiced sonorants. The situation was analogous to the situation in my experiment.

Greenwood tested the learnability of her two languages in a casual speech condition and a careful speech condition. In the casual speech condition, where Greenwood expected the more natural *Z# language to be learned better, there was no difference in how well the two languages were learned. In the careful speech condition, where Greenwood expected the two languages to be learned equally well, the *S# (final voicing) language was learned better than the *Z# (final devoicing) language. Given that the stimuli in my experiment were produced carefully, my study more closely resembles Greenwood's careful speech condition, and our results are analogous. While Greenwood does not consider complexity bias as a possible explanation for her unexpected results, in both our experiments participants learned the training pattern better when they could posit the constraint *[-voice]# than when they had to posit the more complex constraint *[-son, +voice]#. Complexity bias can explain the unusual results of both of our studies.

Although the complexity bias account is promising, its support in my experiment is only partial. The effect of complexity bias only emerged significantly for the initial contrast conditions, #{T, D}…D# and *#{T, D}…T#*. The difference in the acceptance rates of Novel Nonconforming items between these two conditions shows that subjects learned *[-voice] better than *[-son, +voice] when those constraints were against segments in word-final position (*[-voice]# vs. *[-son, +voice]#). When the constraints were against segments in word-initial position (*#[-voice] vs. *#[-son, +voice]), there was a trend for *[-voice] to be learned better than *[-son, +voice], but the difference was not significant. It remains puzzling why the difference in learnability between the simple constraint and the complex constraint should emerge significantly only in word-final position and not in word-initial position.

The present study was designed to investigate substantive bias in phonotactic learning by testing whether learners would reproduce a phonotactic implicational about the existence of an obstruent voicing contrast in different positions in an artificial grammar learning experiment. The phonotactic implicational was not reproduced, and the experiment consequently did not yield any evidence for substantive bias. It did, however, yield some evidence for an effect of complexity bias on phonotactic learning. These results are in line with Moreton & Pater's (2012a,b) conclusion that there is compelling evidence for complexity bias but little for substantive bias in phonological learning. This study provides further support for complexity bias and contributes to the ongoing debate about the degree to which synchronic learning biases, as opposed to diachronic factors, are responsible for the phonological typology.

References

Becker, M., N. Ketrez, & A. Nevins. 2011. The Surfeit of the Stimulus: Analytical Biases Filter Lexical Statistics in Turkish Laryngeal Alternations. *Language* 87.84–125.

Becker, M. & Levine, J. (2013). Experigen – an online experiment platform. http://becker.phonologist.org/experigen.

Becker, M., Nevins, A., & Levine, J. (2012). Asymmetries in Generalizing Alternations To and From Initial Syllables. *Language*, *88*(2), 231–268.

Finley, S. 2012. Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes* 27.1550–1562.

Greenberg, J. H. 1978. Some Generalizations Concerning Initial and Final Consonant Clusters. In *Universals of Human Language, Volume 2, Phonology*, ed. by J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik, 243–280. Stanford: Stanford University Press.

Greenwood, A. 2016. *An Experimental Investigation of Phonetic Naturalness*. Ph.D. dissertation. University of California, Santa Cruz.

Hayes, B., P. Siptár, K. Zuraw, & Z. Londe. 2009. Natural and Unnatural Constraints in Hungarian Vowel Harmony. *Language* 85.822–863.

Hayes, B. & J. White. 2013. Phonological Naturalness and Phonotactic Learning. *Linguistic Inquiry* 44.45–75.

Hayes, B. & C. Wilson. 2008. A Maximum Entropy Model of Phonotactic and Phonotactic Learning. *Linguistic Inquiry* 39.379–440.

Moreton, E. 2008. Analytic bias and phonological typology. *Phonology* 25.83–127.

Moreton, E. 2012. Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language* 67.165–183.

Moreton, E. & J. Pater. 2012a. Structure and Substance in Artificial-phonology Learning, Part I: Structure. *Language and Linguistics Compass* 6.686–701.

Moreton, E. & J. Pater. 2012b. Structure and Substance in Artificial-phonology Learning, Part II: Substance. *Language and Linguistics Compass* 6.702–718.

Myers, S. & J. Padgett. 2014. Domain generalisation in artificial language learning. *Phonology* 31.399–433.

Skoruppa, K. & S. Peperkamp. 2011. Adaptation to Novel Accents: Feature-Based Learning of Context-Sensitive Phonological Regularities. *Cognitive Science* 35.348–366.

Steriade, D. 1987. Redundant Values. In *Proc. of CLS 23: Parasession on Autosegmental and Metrical Phonology*, 339–362.

Steriade, D. 1997. Phonetics in phonology: The case of laryngeal neutralization. Ms. University of California, Los Angeles.

White, J. 2013. *Bias in phonological learning: Evidence from saltation*. Ph.D. dissertation. University of California, Los Angeles.

Wilson, C. 2006. Learning phonology with a substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science* 30.945–982.