2012

Structure and Substance in Artificial-Phonology Learning, Part II: Substance

Elliott Moreton, *University of North Carolina at Chapel Hill*
Joe Pater

Structure and substance in artificial-phonology learning, Part II: Substance

Elliott Moreton — UNC-Chapel Hill
Joe Pater — UMass-Amherst *

Version 5.0, June 29, 2012

Abstract

Artificial analogues of natural-language phonological patterns can often be learned in the lab from small amounts of training or exposure. The difficulty of a featurally-defined pattern has been hypothesized to be affected by two main factors, its formal structure (the abstract logical relationships between the defining features) and its phonetic substance (the concrete phonetic interpretation of the pattern). This paper, the second of a two-part series, reviews the experimental literature on phonetic substance, which is hypothesized to facilitate the acquisition of phonological patterns that resemble naturally-occurring phonetic patterns. The effects of phonetic substance on pattern learning turn out to be elusive and unreliable in comparison with the robust effects of formal complexity (reviewed in Part I). If natural-language acquisition is guided by the same inductive biases as are found in the lab, these results support a theory in which inductive bias shapes only the form, and channel bias shapes the content, of the sound patterns of the worlds languages.

Contents

*The authors gratefully acknowledge the advice and suggestions already contributed by two anonymous reviewers, Bruce Hayes, Anne Pycha, Jen Smith, Kie Zuraw, and the participants in Linguistics 751 at the University of Massachusetts in Spring 2011, especially John Kingston, Claire Moore-Cantwell, Wendell Kimper, and Robert Staubs. This paper is Part II of a two-paper series, and is intended to be read in conjunction with Part I.

1 Introduction

The formal structure of a featurally-defined phonological pattern has repeatedly been found to affect its learnability in the lab. In particular, patterns are harder to learn if they are more complex in certain ways, such as making crucial reference to more feature instances or to more feature types (see our companion paper, Part I, for a review). However, patterns can differ in ways other than the number of features and the logical relationship between them. Perhaps the most important of these is in phonetic substance, i.e., the concrete interpretation of the pattern in terms of phonetic variables. In this paper, we review the evidence for phonetically-sensitive inductive bias in artificial-phonology learning, and contrast it with the evidence for structure-based inductive bias reviewed in Part I.

In natural language, patterns with similar formal structure can have very different typological frequencies depending on how those patterns are instantiated phonetically. For example, coronal-stop assibilation is asymmetrically triggered by following rather than preceding vocoids, and by high rather than low vocoids (Kim, 2001), and vowel-height harmony is more common than consonant continuancy harmony (Hansson 2001, 137–149, Rose and Walker 2004). Very often, the more-frequent phonological pattern is a stylized version of some kind of phonetic covariation (e.g., phonological vowel harmony resembles phonetic vowel-to-vowel coarticulation). One family of explanations proposes that learners have an inductive bias in favor patterns which have phonetic motivations over those which do not (e.g., Stampe, 1973; Archangeli and Pulleyblank, 1994; Steriade, 1997; Hayes, 1999; Steriade, 2001; Wilson, 2006). The experiments that test this hypothesis typically compare a typologically-frequent, phonetically-motivated pattern to a rare, unmotivated (or even counter-motivated) pattern of the same complexity. The available studies can be divided into three categories: consonant/vowel asymmetries, segmental patterns, and prosodic patterns.

2 Consonants versus vowels

In natural language, phonological dependencies between non-adjacent vowels are thought to be much more common than dependencies between non-adjacent consonants (e.g., Gafos (1996, 7–8), Baković (2000, 5–6), Hansson (2001, 1–2)), Attention in the lab has focused on a specific kind of non-adjacent dependency, namely repetition.

There have been several reports that a within-stimulus vowel-repetition dependency is learned faster than the analogous consonant pattern. Toro, Nespor, Mehler, and Bonatti (2008) familiarized native Italian speakers on $CVCVCV$ stimuli. In one condition, the first and last vowel were identical; in the other, the first and last consonant. Participants in the vowel condition preferred novel positive stimuli over non-conforming foils; those in the consonant condition did not (the two conditions were not directly compared to each other). The finding was replicated by Toro, Shukla, Nespor, and Endress (2008), where it was also shown to be robust against manipulation of the relative amplitudes of the consonants and vowels (again without direct comparison). It was replicated by Pons and Toro (2010) with Spanish-learning 16-month-olds; however, a direct comparison between the two conditions found no significant difference. Nevins and Toro (Nevins, 2010), in an experiment with Italian-speaking adults, directly compared consonant- with vowel-repetition patterns, and found a stronger preference for positive instances with the latter.

On the other hand, some studies have found no advantage for vowel repetition over consonant repetition. Koo (2007, Ch. 2), using a speeded-repetition paradigm with English speakers, found a conformity advantage for an [l. . . l]/[r. . . r] pattern and an [l. . . r]/[r. . . l] pattern, but not for analogous patterns with [i] and [u]. Since the responses in this experiment are utterances rather than judgments, the result may be due to the articulatory difficulty of co-occurring liquids, rather than to differences in learnability of the patterns. Two subsequent experiments in the same series with [l. . . l]/[r. . . r] and [i. . . i]/[u. . . u] found no difference in their effects on pattern-membership judgments of new stimuli.

On the whole, the experimental evidence is consistent with the hypothesis that learners are more sensitive to syntagmatic repetition of vowels than of consonants. If there is such an inductive bias, and if that bias shapes typology, we should find that patterns of non-adjacent vowel repetition outnumber the analogous consonant patterns in natural languages. Many languages have a phonological pattern in which one vowel is required to be identical to another, notably total vowel harmony (Aoki, 1968) and copy-vowel epenthesis (Kitto and de Lacy, 1999), whereas the consonantal analogues seem to us to be much rarer. However, we know of no quantitative test of this hypothesis.

Aside from the identity dependency, isomorphic consonant and vowel patterns have seldom been directly compared. Zaba (2008) found no differences between vowel backness agreement and two consonant featural-agreement patterns (see §3). Moreton (2012), using a paradigm described in our companion paper (Part I, §3.1) found a modest but significant advantage for two phonetically-systematic non-adjacent vowel dependencies over two phonetically-systematic non-adjacent consonant dependencies. It is not known whether this difference would persist if other, possibly more salient, consonant features were used instead of place and voicing (LaRiviere et al., 1974, 1977).

3 Segmental harmony

In natural language, dissimilation is thought to be rarer than assimilation both synchronically (Bye, 2011) and diachronically (Campbell, 2004, 30). Some studies have tested corresponding agreement and disagreement patterns side by side, but there is hardly any evidence that agreement is easier to learn. As described in our companion paper (Part I, §3), Pycha et al. (2003) found no difference in learnability between analogues of backness harmony and backness disharmony; Wilson (2003) found no difference between nasal agreement and disagreement; and Kuo 2009, 139 found no difference between place agreement and place disagreement (labial glide iff labial stop, vs. labial glide iff coronal stop). Similar results have been found in other studies as well.

Koo (2007, §2.1; Koo and Cole, 2006) used a speeded-repetition paradigm to familiarize English speakers with long-distance agreement and disagreement patterns for liquid laterality ([l] vs. [ɹ]) and vowel backness ([i e] vs. [a u]). New pattern-conforming test items were repeated quicker than non-conforming foils in both of the liquid conditions, and in neither of the backness conditions. The only sign of a harmonic-disharmonic difference was a higher error rate on non-conforming items in the liquid-agreement condition, alongside no difference in the liquid-disagreemnt condition. (The interpretation of these experiments is complicated somewhat by the absence of a baseline condition in which participants are familiarized on neither pattern. Since we do not know what performance is like for speakers unfamiliar with both pattern, we cannot be sure how much of the performance in the critical conditions is due to learning.)

Finally, Skoruppa and Peperkamp (2011) modified the non-low front vowels in Standard French words to create "Harmonic French", in which the vowels of a word agree in rounding (e.g., *pudeur* [pydœʁ] → [pydɛʁ]), and "Disharmonic French", in which vowels in odd- and even-numbered positions disagree in rounding (e.g., *ordinaire* [ɔʁdinɛʁ] → [ɔʁdinœʁ]). Participants were familiarized by hearing stories in one or the other accent. then tested by hearing paired Harmonic and Disharmonic French versions of words and trying to choose the one in the familiar "accent". This they did better than chance, but it made no significant difference whether the familiarized accent was Harmonic or Disharmonic. Peperkamp and Skoruppa conclude (p. 356) that there are "no abstract linguistic biases favoring harmony over disharmony in perceptual phonological learning"; the preponderance of assimilation over dissimilation in natural language must have other causes.

Other studies have compared harmony patterns differing in typological frequency. Usually, no significant advantage is found for the analogue of the more-common pattern.

In natural languages, rounding harmony in mid vowels asymmetrically implies rounding harmony in high vowels (Kaun, 2004). In an experimental analogue, Finley and Badecker

(2009, Exp. 3) familiarized English speakers, by passive listening, with stimuli of the form $X \ldots XY$ (where X was a $CVCV$ nonsense word and Y a CV suffix), and then asked them to choose the positive member of a pair of new stimuli XZ_1, XZ_2. The vowel of Y agreed in backness and rounding with those of X. For one group of participants, the vowels of Y were high and those of the Z's were mid; the reverse was true for the other group. Although familiarized participants chose the positive stimuli much more often than did unfamiliarized control listeners, the rate did not differ between the two familiarization groups. In other words, no analogue of the asymmetrical implication in natural language typology was found.

Zaba (2008, Ch. 2) used a paradigm similar to that of Pycha et al. (2003) to compare artificial analogues of three patterns differing in natural-language typological frequency: backness agreement between non-adjacent vowels (common), nasality agreement between non-adjacent consonants (rare), and labiality ageement between non-adjacent consonants (unattested in adult language). English-speaking participants were familiarized by listening to stimuli of the form $X \ldots XY$, where X ended in $\ldots V_1C_1$, and Y could be either of two V_2C_2 sequences. The pattern deterimined the choice of Y as a function of V_1 (for backness agreement) or C_1 (for nasality and labiality agreement). Learning was then tested by asking participants to judge whether $X \ldots XY$ sequences conformed to the trained pattern. A control group of participants was likewise tested after being familiarized with $X \ldots XY$ stimuli in which Y was always the same. The results were analyzed in several different ways, none of which found significant differences between the three pattern conditions.

Sibilant harmony in natural languages can be triggered by only [–anterior] segments, only [+anterior] segments, or both. The first of these possibilities is much rarer typologically than the other two. Kosa (2010) compared two artificial analogues of rightward-spreading sibilant-harmony patterns, one triggered by [–anterior] and the other by [+anterior]. Words had the form XY, where X was one of many $CVCV$ sequences. When X contained no sibilants, Y was one of [næ], [sʌt], or [ʃɑ]. For participants in the [–anterior] condition, [sɑ] replaced [ʃɑ] when X contained a [–anterior] sibilant; for those in the [+anterior] condition, [ʃʌt] replaced [sʌt] when X contained a [+anterior] sibilant. In Experiment A, English-speaking participants listened to pattern-conforming XY words and then gave familiar/unfamiliar judgments to conforming and non-conforming test items. Novel items were more likely to be judged familiar when they fit the pattern, but the difference reached significance only in the [+anterior] condition (the analogue of the common natural-language pattern). However, the analysis did not directly compare the two conditions, so we do not know whether they differed significantly from each other. In Experiment B, participants listened to and repeated pattern-conforming XY words, then heard conforming-nonconforming pairs and tried to choose the conforming word. Proportion correct was significantly above chance in both conditions, but did not differ between them.

4 Consonant-vowel interactions

Wilson (2006) focused on two typological asymmetries in rules changing velars [k g] to palatoalveolars [tʃ dʒ] as a function of vowel context. One is that palatalization before more-back vowels implies palatalization before less-back ones; the other, that palatalization of voiced velars implies that of voiceless ones. English speakers were trained in a language game to respond to a subset of [ki gi ke ge] with [tʃi dʒi tʃe dʒe], and to both of [ka ga] with [ka ga] (the critical syllables occurred initially in disyllabic nonsense words). They were then tested on a mix of old and new stimuli to measure their velar-palatalization rate in different conditions. Experiment 1 focused on the effect of vowel context; Experiment 2, that of consonant voicing. A synopsis is given in Table (1).

(1) *Critical experimental conditions of Wilson (2006). n, number of stimuli; p, probability of velar palatalization (in the training stimuli or in the test responses).*

			ki		*ke*		*ka*		*gi*		*ge*		*ga*	
Exp	*Cond*	*Phase*	*n*	*p*	*n*	*p*	*n*	*p*	*n*	*p*	*n*	*p*	*n*	*p*
1	*i*	*Train*	*4*	*1.00*	–	–	*3*	*0.00*	*4*	*1.00*	–	–	*3*	*0.00*
		Test	*8*	*0.44*	*8*	*0.13*	*6*	*0.05*	*8*	*0.52*	*8*	*0.14*	*6*	*0.14*
1	*e*	*Train*	–	–	*4*	*1.00*	*3*	*0.00*	–	–	*4*	*1.00*	*3*	*0.00*
		Test	*8*	*0.20*	*8*	*0.19*	*6*	*0.15*	*8*	*0.48*	*8*	*0.49*	*6*	*0.39*
2	*k*	*Train*	*4*	*1.00*	*4*	*1.00*	*3*	*0.00*	*1*	*1.00*	*1*	*1.00*	*3*	*0.00*
		Test	*8*	*0.39*	*8*	*0.36*	*6*	*0.12*	*8*	*0.14*	*8*	*0.11*	*6*	*0.09*
2	*g*	*Train*	*1*	*1.00*	*1*	*1.00*	*3*	*0.00*	*4*	*1.00*	*4*	*1.00*	*3*	*0.00*
		Test	*8*	*0.26*	*8*	*0.20*	*6*	*0.00*	*8*	*0.50*	*8*	*0.44*	*6*	*0.23*

The rate of palatoalveolar responses to velar-initial stimuli was bimodal, with clusters around 15% (Low) and 45% (Medium), rates comparable to those found by Peperkamp et al. (2006). Changing the features made a clear difference in performance in Experiment 1. Participants trained to respond [tʃe dʒe ka ga] (1e) had the Medium rate of palatoalveolar responses to [gi ge ga], but the Low rate to [ki ke ka]. Participants trained on [tʃi dʒi ka ga] (1i) had the Medium rate only on [ki] and [gi] themselves. These results are to some extent consistent with typology, since velar palatalization before [e] asymmetrically implies velar palatalization before [i].

However, there are also differences. Participants in Condition 1e, unlike natural-language palatalization rules, disregarded vowel context entirely. Their palatalization rates were indistinguishable before [e], where they had been trained to palatalize, [a], where they had been trained *not* to palatalize, and [i], where they had been given no training (Wilson, *op. cit.*, Fig. 2). In natural language, [g]-palatalization asymmetrically implies [k]-palatalization. If this is a result of inductive bias, we expect the learner to interpret observed [g]-palatalization as evidence for [k]-palatalization, but not the other way around.

This was not borne out in the experiments. In Experiment 1, [g] was palatalized significantly more often than [k], despite equal training on both. In Experiment 2, there was no significant difference between the effects of [k]-training on [g]-palatalization and that of [g]-training on [k]-palatalization.

5 Prosodic rules

Carpenter (2005, 2006, 2010) investigated acquisition of artificial stress patterns by native speakers of American English and Quebec French. Participants were trained to choose between correctly- and incorrectly-stressed versions of the same word, then tested with novel words. One set of experiments used sonority-sensitive stress, comparing the typologically uncommon "leftmost low vowel ([æ a]), else leftmost vowel" with the completely unattested "leftmost high vowel ([i u]), else leftmost vowel". Participants in both native-language groups preferred positive items more strongly in the leftmost-low condition. Here the lab results are aligned with typology.[1]

A second set of experiments in this series used quantity-sensitive stress, comparing the typologically frequent "leftmost heavy (CVC), else leftmost" pattern with the unattested "leftmost light (CV), else leftmost". Since the typology is much more strongly skewed here than in the case of sonority-sensitve stress, one would expect the same of learning performance. However, although preference for positive stimuli was significantly above chance in all conditions, it did not differ significantly between the two artificial patterns in either native-language group.

Stress in natural language is influenced much more often by codas than by onsets. In an experiment with 9-month-old English-learning infants, Gerken and Bollt (2008, Experiment 2) used a familiarization set of polysyllabic pseudowords in which stress was determined in part by the form of the syllables and in part by other principles. The initial $CV(:)$ syllable was always stressed.[2] In one condition, stress preferentially fell on syllables of the form $CV:C$ (e.g., /boʊm/) rather than those of the form $CV:$ (e.g., /mi/) or CV (e.g., /fa/). In the other, stress preferentially fell on syllables that began with /t/ rather than any of /d m f s l ɹ/. Test items always followed both of these form-based generalizations by stressing the unique [t]$V:C$ syllable, but, unlike familiarization items, never stressed the initial $CV(:)$. In half of the test items, these two critical syllables were adjacent and could not both be stressed without clashing. For the other half, both critical syllables could have been stressed without clash (but were not). The rationale is that if participants

[1] Although the vowels were equalized for duration and peak intensity to remove phonetic stress cues, the author cautions us that participants may have restored low vowels' inherently greater duration and intensity by subvocalizing the stimuli (Carpenter, 2010). However, the unedited high vowels were tense, while the unedited low ones were lax. Editing lengthened high vowels and shortened low ones (Carpenter, 2006, 72), exaggerating the tense/lax contrast in a way that would encourage English speakers to stress the high vowels, not the low ones.

[2] In another version of the experiment, initial and final syllables traded roles.

learn the form-based generalization, they will, in effect, realize that test items of the first type have a valid excuse for not stressing the initial syllable, and be more likely to accept them as pattern-conforming. Infants in both conditions listened longer to the test items of the first kind than the second, but the difference only reached significance in the condition where stress was coda-sensitive.

We are reluctant to accept these results as evidence of an inductive bias in favor of coda-sensitive over onset-sensitive stress. First, a direct comparison between the onset and coda conditions (p. 241) found no significant differences between them (i.e., performance in one condition differed significantly from chance, while that in the other did not, but the two conditions did not differ significantly from each other). Second, the two patterns were not directly comparable. The coda-sensitive pattern depended on the presence vs. absence of a consonant, while the onset-sensitive one depended on the identity of an ever-present consonant. The former kind of difference may be more salient to infant pattern learners (Saffran and Thiessen, 2003).

Schane et al. (1974) trained participants to translate English adjective-plus-noun phrases word for word into an artificial language which had a context-sensitive rule deleting the final consonant of the first word (Figure 2). In one condition, deletion applied when the second word began with a consonant, simulating a cross-linguistically common process of cluster simplification (Wilson, 2001). Deletion in the other condition occurred when the second word was vowel-initial. Such intervocalic deletion in nature rarely applies to the voiceless obstruents used in this experiment (Picard, 2003).

(2) *Artificial-language conditions used by Schane et al. (1974).*

Condition	$/C_1\#C_2/$ */ˈtupak ˈsipu/*	$/C_1\#V_2/$ */ˈtupak ˈoga/*
Cluster simplification	$\rightarrow [\#C_2]$ *[ˈtupa ˈsipu]*	*[ˈtupak ˈoga]*
Intervocalic C deletion	*[ˈtupak ˈsipu]*	$\rightarrow [\#V_2]$ *[ˈtupa ˈoga]*

Participants were trained to a performance criterion. The cluster-simplification groups reached criterion before the intervocalic-deletion ones, and were less likely to erroneously give responses appropriate to the other group.

In both conditions, the presence of a consonant at the end of Word 1 was correlated with the presence of a consonant at the beginning of Word 2. Why would the negative correlation be easier to learn than the positive one? One possibility is substantive bias against marked phonological structures such as hiatus or syllable codas. Alternatively, participants in the cluster-simplification condition could have learned the pattern “exactly one intervocalic consonant”, while those in the intervocalic-deletion condition would have had to learn the disjunction “zero or two intervocalic consonants”. The results could then be accounted for by a complexity bias, the relative difficulty of disjunctive categories (Bruner et al., 1956; Conant and Trabasso, 1964; Ciborowski and Cole, 1972).

6 Summary: Substance

The best-supported finding, replicated with different stimuli and with speakers of different ages and languages, that it is easier to induce the pattern "contains the same vowel twice" than "contains the same consonant twice". The corresponding natural-language phenomena are typologically small potatoes. Meanwhile, the clearest non-finding is that, despite several attempts motivated by the typological importance of the patterns, no study has yet found positive evidence that vowel or consonant harmony is easier to learn than disharmony.

For other segmentally- and prosodically-defined patterns, results have been mixed at best, even within the same study. Even when positive (statistically significant) results are found, they may fail to match typology. For example, [g] was palatalized more than [k] in Wilson (2006, Experiment 1), whereas in natural language, [g]-palatalization asymmetrically implies [k]-palatalization. Such mismatches are hard to account for. When complexity bias mismatches typological frequency, the disparity may be due to the competing effects of channel bias on typology (Moreton, 2008). This explanation is not available when substantive bias mismatches typology, since channel bias is predicted to *reinforce* substantive bias in shaping typology.

7 General discussion

An isolated "null result" (lack of significant difference between conditions) is hardly fatal to a hypothesis. It could be due to sampling error, or a flawed experiment. However, the null results in substantive-bias experiments are not isolated; they contrast with the positive results routinely obtained for complexity differences using similar experimental paradigms and participant populations. In our companion paper (Part I, §2), for example, we saw that significant advantages for featurally-simpler patterns were found by Pycha et al. (2003); Saffran and Thiessen (2003); Peperkamp et al. (2006); Cristiá and Seidl (2008); Kuo (2009); Skoruppa et al. (2009); Chambers et al. (2010); Skoruppa and Peperkamp (2011). Only one study yielded ambiguous results: LaRiviere et al. (1974, 1977) found a numerical advantage for the simpler pattern in ten experiments, but it only reached statistical significance in three of them. (The complexity affect vanishes in younger infants; Cristià, Seidl, and Gerken, 2011; Cristià, Seidl, and Francis, 2011). On the other hand, when experimenters compared phonetically-motived harmony patterns with corresponding disharmony patterns (§3), the opposite occurred: Not one study found an advantage for the harmony pattern (Pycha et al., 2003; Wilson, 2003; Koo and Cole, 2006; Kuo, 2009; Skoruppa and Peperkamp, 2011). Each individual null result could be attributed to chance, but the concentration of null results on the substance side of the ledger cannot be.[3]

The lab results so far thus add up to substantial and, in our view, convincing evidence that formal complexity impedes artificial-phonology learning in adults and infants, but the effect of phonetic substance is weaker if it exists at all. These findings corroborate the suggestion of Pycha et al. (2003) that complexity bias is stronger than substantive bias in artificial phonology. The leap from the existing artificial-phonology evidence to firm conclusions about natural-language phonology is a big one. It is decidedly better than a leap from no evidence at all, but much work needs to be done before the laboratory study of inductive biases approaches the level of reliability and sophistication found in that of channel biases. In the following sections, we consider some of the theoretical, empirical, and practical issues involved, and offer some suggestions for narrowing the gap.

7.1 An analytic artifact?

Calculations of both complexity and substance are affected by the analyst's choice of featural primitives, and phonological-feature models are deliberately engineered to encode substantive bias as complexity bias (McCarthy, 1988; Clements and Hume, 1995). This theoretical prejudice could inflate the apparent frequency of complexity biases at the expense of substantive ones. For example, when Cristiá and Seidl (2008) found that [m n b k] vs. [f z] was easier to learn than [m n f z] vs. [b k], we interpreted it as evidence that Type I patterns are easier than Type II (see Part I, §2). That interpretation depended on the absence of a phonological feature distinguishing [m n f z] from [b k]. If the learner in fact has such a feature (e.g., [continuous airflow] or [stable spectrum]), then reluctance to use it would constitute a substantive bias against a Type I pattern, rather than a structural bias against a Type II pattern.

Other experiments, however, resist such reanalysis. To convert the Saffran and Thiessen (2003) experiment from Type I vs. Type II into Type I vs. Type I, we would have to credit the learner with a phonological feature that separates [p d k] from [b t g]. To convert Kuo (2009)'s Type II vs. Type VI into Type II vs. Type II would require one that separates [p^h t] from [p t^h], and Pycha et al. (2003)'s II/VI comparison would need one that separated [i æ ʊ] from [u a ɪ]. Such phonetically arbitrary, *post hoc* features cannot be universal, nor learned from prior experience; they could only have been learned in the experiment itself. In that case, the unifying principle behind the experimental results would be that a pattern is harder when a phonetically complex feature has to be induced — a structural bias, not a substantive one.

The substantive-bias hypothesis is not just about classes of sounds, but about their behavior. A pattern does not qualify as phonetically motivated merely by employing a

[3]If we count the studies listed in this paragraph as 8 successes and 1 failure among the complexity studies, and no successes and 5 failures among the substance studies, then the difference in success rate is significant by Fisher's exact test ($p < 0.003$). A more sophisticated statistical meta-analysis would compare individual experiments, taking into account properties shared between experiments.

phonetically-defined segment class; that class has to *do* something for which there is a phonetic explanation. For example, final-obstruent devoicing and final-obstruent voicing both involve obstruents, but only the former is phonetically motivated (Steriade, 1997; Kiparsky, 2008). Although the Type I patterns used by LaRiviere et al. (1974, 1977); Saffran and Thiessen (2003), and Cristiá and Seidl (2008) were built on typologically common sound classes, the patterns themselves were phonetically unmotivated and typologically rare (e.g., "onsets must be nasal or oral stops"). Phonetic motivation therefore does not explain why they were learned better than equally unmotivated Type II patterns. Pycha et al. (2003) found no statistically significant difference in difficulty between a phonetically motivated, typologically common Type II pattern of vowel backness agreement and a phonetically less-motivated, typologically rare Type II pattern of backness disagreement, though both proved easier than a Type VI pattern. The three Type II patterns compared by Zaba (2008, Ch. 2) differed widely in phonetic motivation and typological frequency, but not in learning outcome.

7.2 Is artificial phonology phonology at all?

Perhaps participants are treating the artificial-phonology task as if it were a non-linguistic concept-learning task about red triangles or fictitious diseases. Use of the same domain-general cognitive processes would predict shared complexity biases and lack of substantive bias. It may be that such domain-general mechanisms are involved in natural-language phonology as well (Hume and Johnson, 2001). If so, then artificial phonology is informative about natural-language phonology after all. However, it is also possible that natural-language phonology is learned using a separate set of dedicated processes, and hence that artificial phonology is irrelevant to it. Several strands of evidence are germane to this hypothesis.

The sharing of structural biases between artificial phonology and non-linguistic category learning does not prove that they share a common mechanism to the exclusion of natural-language phonology. The currently known shared biases are few, and most are so generic that even radically different learning algorithms can share them (Gluck and Bower, 1988; Kruschke, 1992; Nosofsky et al., 1994; Love et al., 2004; Feldman, 2006). Nature is not guaranteed to be parsimonious, sometimes building two physically distinct circuits where behavioral similarity would lead us to expect one (e.g., Simons-Wiedenmaier et al., 2006). Furthermore, artificial phonology also shares structural biases with natural-language phonological typology (see Part I, §4).

Since substantive biases are by nature domain-specific, they could provide very strong evidence that artificial phonology is learned using cognitive processes peculiar to phonology — which could only be the same processes used for learning natural-language phonology. The more arbitrary and stipulative the bias, the stronger would be the evidence that artificial phonology experiments are directly relevant to natural language. However, as we have seen, the evidence for substantive biases is weak.

The separate-mechanism hypothesis would also be strengthened if there were an age at which infants show no knowledge of first-language phonology, but can rapidly learn artificial phonology. Like adults, infants receive much more exposure to the ambient language than to any experimental one. If artificial and natural phonology are acquired by the same mechanisms, then anyone old enough to be influenced by a few minutes' exposure to an artificial pattern should also be old enough to be influenced by comparable patterns in the phonology of the ambient language (where "comparable" means that the patterns have similar complexity and are instantiated with equal consistency in the respective training-data sets). For example, by six months of age, Turkish-learning infants, unlike German-learning ones, spontaneously prefer disyllabic pseudowords which have backness harmony over those which lack it (van Kampen et al., 2008). If Turkish-learning 4.5-month-olds do not have this preference spontaneously, but do acquire it from brief exposure to Turkish-like stimuli in an artificial-language experiment, that would support the separate-mechanism hypothesis.

Infants start showing sensitivity to first-language phonotactic patterns between 6 and 9 months of age (Jusczyk et al., 1993; Friederici and Wessels, 1993; Jusczyk et al., 1994; Mattys and Jusczyk, 2001). Cristià and colleagues (Cristià, Seidl, and Gerken, 2011; Cristià, Seidl, and Francis, 2011), using a paradigm described in our companion paper (Part I, §2), found that infants as young as 4 months acquired a novelty preference after a few minutes' experience with a positional restriction on consonant occurrence, succeeding on both a Type I and a Type II pattern where 7-month-olds failed on the Type II pattern. Reviewing a range of studies, they conclude that complexity effects begin to emerge at around 7 to 9 months.

Interactions between natural and artificial phonology are evidence that the two are not wholly separate processes. Schane et al. (1974)'s cluster-simplification pattern was, by design, similar to French liaison, and was acquired better by those who had had more exposure to French. Healy and Levitt (1980, Experiment 2) found better performance with a pattern that resembled English voicing agreement (in -*ed*, -*'s*, etc.) than for one that did not. Pater and Tessier (2006), again with English speakers, found better performance on an artificial epenthesis rule when it was triggered by violation of the English minimal-word constraint than when it was triggered by vowel backness. It is not clear whether natural language phonology had these effects by facilitating rule learning in training, or by biasing participants' responses against non-conforming responses in testing, but it is clear that some aspect of the artificial-phonology task are ecologically valid enough to engage natural-language phonological preferences.

7.3 Exposure and robustness

An obvious difference between artificial-phonology experiments and natural first- or second-language phonology is amount and duration of exposure. Most of the studies reviewed above used less than 30 minutes of familiarization or training, and some used much less.

Taylor and Houghton (2005, Exp. 5) used a tongue-twister paradigm to familiarize English speakers with a pattern that restricted different consonants to different syllabic positions (onset vs. coda). They then reversed the pattern unannounced in the middle of a block. Before the switch, transposition errors adhered to the original constraints 98% of the time. After it, the new constraints were adhered to 65% of the time, with the violations concentrated in the first 9 trials after the switch. In contrast, first-language phonology is notoriously persistent, even in highly motivated speakers with ample exposure to second-language input (Cutler et al., 1989; Darcy, 2006; Kager et al., 2008).[4]

Artificial phonology thus appears easier to lose than first-language phonology and easier to acquire than second-language phonology, which seems to support the different-mechanism hypothesis. However, the difference may not be between a linguistic and a non-linguistic mechanism, but between a short-term and a long-term one. In most artificial-phonology experiments, participants are tested immediately after training, which gives only those learning processes that are active during training an opportunity to apply. Learning actually continues long after the stimuli are experienced, as memories are consolidated and reprocessed. Several studies have found increased sensitivity to patterns when the interval between training and test includes sleep compared to when it does not (Wagner et al., 2004; Gómez et al., 2006; St. Clair and Monaghan, 2008). Integration of new words into the lexicon likewise continues after exposure (see Lindsay and Gaskell 2010 for a recent review), perhaps affecting the time course of their availability to pattern-finding processes. Artificial-phonology experiments using longer time scales may be therefore able to detect subtler effects than the ones reviewed above.

7.4 Structurally biased phonology

In short-term artificial-phonology learning, the effects of phonetic substance, if they exist at all, are overshadowed by those of formal complexity. If natural-language phonological learning works the same way, then the broader picture is essentially that proposed by Bach and Harms (1972): Inductive bias, a property of the learner's pattern-detection processes, facilitates faithful acquisition of simple patterns and rejection or innovative simplification of more complex ones, but is (relatively) insensitive to their phonetic motivation. Channel bias, a property of the articulatory-acoustic-perceptual channel, systematically distorts the phonological form of utterances in transmission, introducing new patterns into the learner's input (Hyman, 1976; Ohala, 1992, 1993). Together, the complexity-based inductive biases and the phonetic channel biases cause systematic asymmetries in the direction of language change, and hence in the long-term steady-state typological frequencies of different kinds

[4] *Some* first-language patterns can be overcome with a small amount of lab exposure. Carpenter (2005, 2006, 2010) succeeded in teaching English speakers the artificial stress rule "stress leftmost low vowel, else leftmost vowel". Since the low vowels were also lax and the high ones also tense, the artificial pattern overcame the English L1 preference for stressing heavy syllables. Kuo (2009) found that Mandarin speakers' L1 phonotactics did not affect their ability to acquire L1-conforming vs. L1-nonconforming artificial patterns from brief exposure.

of pattern (Bell, 1970, 1971; Greenberg, 1978).[5] Over time, phonological patterns are predicted to become featurally simpler and less phonetically-motivated (Bach and Harms, 1972).

This view could be called *Structurally-Biased Phonology*. Its main feature is that it separates phonetic from structural factors, identifying the former with channel bias and the latter with inductive bias, while abolishing the distinction between "synchronic" and "diachronic" factors — both kinds of bias are observable synchronically, in the lab and in their natural setting, and both shape typology only by skewing the outcome of diachronic change.[6] Elaborating and testing this hypothesis raises several research questions, a few of which we will touch on here (see Pater and Moreton (in press) for expanded discussion).

What are the inductive biases? What do they reveal about the learner? The work reviewed in this paper is a good start, but much still needs doing to perfect methodologies for studying inductive biases, to discover what kinds of pattern they favor, and to determine what properties a learner would have to have in order to exhibit them. Since a major source of inductive bias is the architecture of the learner, this mission will necessarily involve contact with models of non-linguistic pattern learning in psychology and computer science. Another major source of inductive bias is the phonological-representation schema used by the learner, since complexity is only meaningful in terms of the representational primitives available for expressing generalizations. This was the central insight behind Feature Geometry, a framework in which the feature system was engineered to allow all and only typologically common, phonetically transparent processes to be expressed by elementary operations of spreading and delinking (Clements 1985; McCarthy 1988; Clements and Hume 1995; for a critical review, see Padgett 2002a, 2002b), while more complex processes required unlikely rule conspiracies (Pulleyblank, 1988, 299). Structurally-Biased Phonology requires a renewed focus on feature systems, and provides a new tool for assessing them: their ability to predict complexity effects in the lab and in typology. (For example, the Cristiá and Seidl (2008) experiment, as discussed in §7.1 is evidence that human infants do not analyze patterns in terms of [continuous airflow] or [stable spectrum].)

What are the channel biases? The richness of the phonetic literature and the concreteness of the variables creates the impression of solidity: In using channel bias to explain phonological typology, we seem to be using the known to illuminate the unknown. In fact, we still know next to nothing about the relative magnitudes of different channel biases across languages. A channel bias is the probability of a particular kind of misunderstanding between speaker and hearer; e.g., the probability that the listener will hear as harmonic

[5] Social factors are crucial to the initiation, continuation, and spread of change (Weinreich et al., 1968), but we expect them to be neutral as to the phonetic or structural content of the patterns. The typological preponderance of vowel harmony over vowel disharmony, for example, does not have a social explanation. The nonexistence of substantive inductive bias has of course been proposed before, e.g., Ohala (1997); Hale and Reiss (2000); Hume and Johnson (2001).

[6] Structurally-Biased Phonology differs from Kiparsky's (2006, 2008) Amphichronic Program in this separation of phonetic from structural factors.

a vowel sequence intended by the listener to be disharmonic. In order to make a prediction about phonological typology, such as whether vowel height harmony should be more frequent across languages than consonant voicing harmony, we need to know at the very least which of them is more strongly favored by channel bias across languages. The data needed to do this does not yet exist. The most convenient point of access to the speaker-hearer channel is its acoustic stage, and in a few cases there are pairs of phonetic dimensions whose interaction has been measured acoustically in a relatively large number of languages; for instance, vowel height and vowel f_0 Whalen and Levitt (1995), or consonant voicing and vowel f_0 (reviewed in Moreton (2010)[7] However, these studies are few, and, as Yu (2012) notes, they only measure acoustic effects, whereas channel bias is hypothesized to affect typology via mis*perception*. Nor do acoustic measurements allow us to compare the magnitude of channel biases affecting different phonetic dimensions, such as height harmony and voicing harmony. This lack of knowledge should not be a deterrent, but a stimulus, to the rigorous testing of hypotheses about channel biases through the study of phonetic typology.

How complex are channel biases? Much previous research has asked whether inductive biases favor phonetically-motivated patterns. Once we turn our attention to explaining the featural complexity of natural-language patterns, we have to ask the complementary question: Do channel biases favor featurally simple patterns? To illustrate the point that inductive bias simplifies complex phonetic relationships, Hayes (1999) demonstrates the complex relationship between the difficulty of stop-closure voicing on the one hand and place of articulation, closure duration, phrasal position, and adjacent nasality on the other.[8] Is that complexity typical of phonetic interactions in general, or can they also be a source of simpicity?

What is the predicted typology? One straightforward prediction is that, other things being equal, if two phonological patterns have the same structure, the typologically more-frequent one should be favored by channel bias, whereas if they are equally favored by channel bias, the more frequent one should be favored by inductive bias (Moreton, 2008). However, since phonological typology is hypothesized to be the long-term steady-state outcome of a Markov process, its relationship to the magnitudes of channel and inductive biases may not be transparent. Something which is unlikely to be innovated can become very frequent if it is even less likely to be extinguished, for example. We expect that in many cases iterated-learning simulations will be needed to derive predictions of the steady-state frequencies. These may also yield testable predictions about the frequency

[7]Even these studies did not really measure the interaction between two phonetic dimensions. Instead, the independent variable was actually phonological, e.g., Whalen and Levitt (1995) looked at how vowel f_0 was affected by vowel height, not by vowel F_1.

[8]Production difficulty is not itself a channel bias, since it is not a probability of an error in transmission between speaker and hearer. However, difficult productions are likely to be produced wrong (that is what it means to say that they are difficult), and hence likely to lead to transmission errors. Consequently, production difficulty is probably positively correlated with channel bias.

and direction of historical change.

What is the actual typology of structural complexity? Phonological typology is usually stated in substantive terms, classifying phonological patterns as vowel harmony, nasal place assimilation, etc. Structurally Biased Phonology makes predictions about the typology of structure and complexity, which has been little studied (Mielke, 2004, e.g.,). How frequent are, for instance, the different Shepard types in the phonological patterns of the world? And do those frequencies differ from those we would expect in the absence of specific complexity biases in the learner? This question leads to the theoretical problem of what the appropriate chance model is (Mackie and Mielke, 2011).

How does structure affect knowledge of natural language? In natural language, Structurally-Biased Phonology predicts that less-complex patterns will be learned better than more-complex ones, and that structurally ambiguous patterns will tend to be generalized in simple ways rather than complex ones. In "poverty-of-the-stimulus" experiments, the experimenter chooses an aspect of the native language phonology which is consistent with two different generalizations, then tests participants' reactions to novel stimulus types to discover which pattern they have actually acquired and infer the inductive bias that guided that acquisition. Complementary "surfeit-of-the-stimulus" (Becker et al., 2007, 2011) experiments compare the productivity of two patterns in the language. Experiments of this sort have been focused on substantive bias (Pertz and Bever, 1975; Davidson et al., 2004; Davidson, 2006; Kawahara, 2006; Zhang et al., 2006; Becker et al., 2007; Zuraw, 2007; Berent et al., 2007, 2008; Moreton et al., 2008; Berent et al., 2009; Hayes et al., 2009; Zhang and Lai, 2010; Becker et al., 2011; Daland et al., 2011; Zhang et al., 2011), but seeing how complexity affects the acquisition of artificial phonology, one would expect strong effects in first- and second-language acquisition as well.

Theories of phonological typology tend to rely primarily on one explanatory factor, channel bias or inductive bias, despite outside evidence that both kinds of bias exist. This reliance may be motivated by regard for theoretical parsimony, since the number of two-factor hypotheses seems unmanageably larger than the already large number of single-factor hypotheses. The main theoretical significance of the evidence we have reviewed in this article is that this dilemma may be avoidable. What Structurally-Biased Phonology offers is a way to control the combinatorial explosion while recognizing the existence and typological effectiveness of both factors.

References

Aoki, H. (1968). Towards a typology of vowel harmony. *International Journal of American Linguistics 34*(2), 142–145.

Archangeli, D. and D. Pulleyblank (1994). *Grounded Phonology.* Cambridge, Massachusetts: MIT Press.

Bach, E. and R. T. Harms (1972). How do languages get crazy rules? In R. P. Stockwell and R. K. S. Macaulay (Eds.), *Linguistic change and generative theory*, Chapter 1, pp. 1–21. Bloomington: Indiana University Press.

Baković, E. (2000). *Harmony, dominance, and control.* Ph. D. thesis, Rutgers University, New Brunswick, New Jersey.

Becker, M., N. Ketrez, and A. Nevins (2007). Where and why to ignore lexical patterns in Turkish obstruent alternations. Handout, 81st annual meeting of the Linguistic Society of America, Anaheim, California.

Becker, M., N. Ketrez, and A. Nevins (2011). The surfeit of the stimulus: analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language 87*(1), 84–125.

Bell, A. (1970). *A state-process approach to syllabicity and syllabic structure.* Ph. D. thesis, Stanford University.

Bell, A. (1971). Some patterns of the occurrence and formation of syllabic structure. *Working Papers on Language Universals 6*, 23–138.

Berent, I., T. Lennertz, J. Jun, M. A. Moreno, and P. Smolensky (2008). Language universals in human brains. *Proceedings of the National Academy of Sciences, U.S.A. 105*(14), 5321–5325.

Berent, I., T. Lennertz, P. Smolensky, and V. Vaknin-Nusbaum (2009). Listeners' knowledge of phonologial universals: evidence from nasal clusters. *Phonology 26*(1), 75–108.

Berent, I., D. Steriade, T. Lennertz, and V. Vaknin (2007). What we know about what we have never heard: evidence from perceptual illusions. *Cognition 104*, 591–630.

Bruner, J. S., J. J. Goodnow, and G. A. Austin (1956). *A study of thinking.* New York: John Wiley and Sons.

Bye, P. (2011). Dissimilation. In M. van Oostendorp, C. Ewen, and E. Hume (Eds.), *The Blackwell companion to phonology*, Chapter 63. Wiley-Blackwell.

Campbell, L. (2004). *Historical linguistics: an introduction* (2nd ed.). Cambridge, Massachusetts: MIT Press.

Carpenter, A. C. (2005). Acquisition of a natural vs. an unnatural stress system. In A. Brugos, M. R. Clark-Cotton, and S. Han (Eds.), *Papers from the 29th Boston University Conference on Language Development (BUCLD 29)*, Somerville, pp. 134–143. Cascadilla Press.

Carpenter, A. C. (2006). *Acquisition of a natural versus an unnatural stress system.* Ph. D. thesis, University of Massachusetts, Amherst.

Carpenter, A. C. (2010). A naturalness bias in learning stress. *Phonology 27*(3), 345–392.

Chambers, K. E., K. H. Onishi, and C. Fisher (2010). A vowel is a vowel: generalizing newly learned phonotactic constraints to new contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition 36*(3), 821–828.

Ciborowski, T. and M. Cole (1972). A cross-cultural study of conjunctive and disjunctive concept learning. *Child Development 43*, 774–789.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook 2*, 225–252.

Clements, G. N. and E. V. Hume (1995). The internal organization of speech sounds. In

J. A. Goldsmith (Ed.), *The handbook of phonological theory*, Chapter 7, pp. 245–306. Boston: Blackwell.

Conant, M. B. and T. Trabasso (1964). Conjunctive and disjunctive concept formation under equal-information conditions. *Journal of Experimental Psychology 67*(3), 250–255.

Cristiá, A. and A. Seidl (2008). Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development 4*(3), 203–227.

Cristià, A., A. Seidl, and A. Francis (2011). Phonological features in infancy. In G. N. Clements and R. Ridouane (Eds.), *Where do phonological contrasts come from? Cognitive, physical, and developmental bases of phonological features*, pp. 303–326. Amsterdam and Philadelphia: John Benjamins.

Cristià, A., A. Seidl, and L. Gerken (2011). Learning classes of sounds in infancy. *University of Pennsylvania Working Papers in Linguistics 17*(1).

Cutler, A., J. Mehler, D. Norris, and J. Segui (1989). Limits on bilingualism. *Nature 340*, 229–230.

Daland, R., B. Hayes, J. White, M. Garellek, A. Davis, and I. Norrmann (2011). Explaining sonority projection effects. *Phonology 29*, 197–234.

Darcy, I. (2006). *Assimilation phonologique et reconnaissance des mots.* Publications Universitaires Européennes, Série XXI: Linguistique. Bern: Peter Lang.

Davidson, L. (2006). Phonology, phonetics, or frequency: influences on the production of non-native sequences. *Journal of Phonetics 34*(1), 104–137.

Davidson, L., P. W. Jusczyk, and P. Smolensky (2004). The initial and final states: theoretical implications and experimental explorations of Richness of the Base. In R. Kager, J. Pater, and W. Zonneveld (Eds.), *Constraints in phonological acquisition*, Chapter 10, pp. 321–368. Cambridge, England: Cambridge University Press.

Feldman, J. (2006). An algebra of human concept learning. *Journal of mathematical psychology 50*, 339–368.

Finley, S. and W. Badecker (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language 61*, 423–437.

Friederici, A. D. and J. M. I. Wessels (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics 54*, 287–295.

Gafos, A. (1996). *The articulatory basis of locality in phonology.* Ph. D. thesis, Johns Hopkins University, Baltimore.

Gerken, L. and A. Bollt (2008). Three exemplars allow at least some linguistic generalizations: implications for generalization mechanisms and constraints. *Language Learning and Development 4*(3), 228–248.

Gluck, M. A. and G. H. Bower (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language 27*, 166–195.

Gómez, R. L., R. R. Bootzin, and L. Nadel (2006). Naps promote abstraction in language-learning infants. *Psychological Science 17*(8), 670–674.

Greenberg, J. H. (1978). Diachrony, synchrony, and language universals. In J. H. Greenberg,

C. A. Ferguson, and E. A. Moravcsik (Eds.), *Universals of human language, volume 1, method and theory*, pp. 61–91. Stanford, California: Stanford University Press.

Hale, M. and C. A. Reiss (2000). 'Substance abuse' and 'dysfunctionalism': current trends in phonology. *Linguistic Inquiry 31*(1), 157–169.

Hansson, G. Ó. (2001). *Theoretical and typological issues in consonant harmony.* Ph. D. thesis, University of California, Berkeley.

Hayes, B. (1999). Phonetically driven phonology: the role of optimality in inductive grounding. In M. Darnell, E. Moravcsik, M. Noonan, F. Newmeyer, and K. Wheatly (Eds.), *Functionalism and Formalism in Linguistics*, Volume 1: General Papers, pp. 243–285. Amsterdam: John Benjamins.

Hayes, B., K. Zuraw, P. Siptár, and Z. Londe (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language 85*(4), 822–863.

Healy, A. F. and A. G. Levitt (1980). Accessibility of the voicing distinction for learning phonological rules. *Memory and Cognition 8*(2), 107–114.

Hume, E. and K. Johnson (2001). A model of the interplay of speech perception and phonology. In E. Hume and K. Johnson (Eds.), *The role of speech perception in phonology*, pp. 3–26. San Diego: Academic Press.

Hyman, L. M. (1976). Phonologization. In A. Juilland (Ed.), *Linguistic studies offered to Joseph Greenberg: second volume: phonology*, pp. 407–418. Saratoga, California: Anma Libri.

Jusczyk, P. W., A. D. Friederici, J. M. I. Wessels, V. Y. Svenkerud, and A. M. Jusczyk (1993). Infants' sensitivity to the sound pattern of native language words. *Journal of Memory and Language 32*, 402–420.

Jusczyk, P. W., P. A. Luce, and J. Charles-Luce (1994). Infants sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language 33*, 630–645.

Kager, R., N. Boll-Avetisyan, and C. Ao (2008, November). Gradient phonotactic constraints for speech segmentation in a second language. Poster presented at the Boston University Conference on Language Development.

Kaun, A. (2004). The typology of rounding harmony. In B. Hayes, R. Kirchner, and D. Steriade (Eds.), *Phonetically based phonology*, Chapter 4, pp. 87–116. Cambridge, England: Cambridge University Press.

Kawahara, S. (2006). Mimetic gemination in Japanese: a challenge for Evolutionary Phonology. *Theoretical Linguistics 32*(3), 411–424.

Kim, H. (2001). A phonetically based account of phonological stop assibilation. *Phonology 18*(1), 81–108.

Kiparsky, P. (2006). The amphichronic program vs. Evolutionary Phonology. *Theoretical Linguistics 32*(2), 217–236.

Kiparsky, P. (2008). Universals constrain change, change results in typological generalizations. In J. Good (Ed.), *Linguistic universals and language change*, Chapter 2, pp. 23–53. Oxford, England: Oxford University Press.

Kitto, C. and P. de Lacy (1999). A Correspondence theory of epenthetic quality. In C. Kitto and C. Smallwood (Eds.), *Proceedings of AFLA (Austronesian Formal Linguistics Association VI*, Toronto Working Papers in Linguistics, pp. 181–200. Toronto: Unspecified publisher.

Koo, H. (2007). *Change in the adult phonological processing system by learning non-adjacent phonotactic constraints from brief experience: an experimental and computational study.* Ph. D. thesis, University of Illinois, Urbana-Champaign.

Koo, H. and J. S. Cole (2006). On learnability and naturalness as constraints on phonological grammar. In *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics, Athens, Greece, August 28-30.*

Kosa, L. A. (2010). Sibilant harmony: investigating the connection between typology and learnability. In *Proceedings of the 2010 annual conference of the Canadian Linguistic Association.*

Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review 99*, 22–44.

Kuo, L. (2009). The role of natural class features in the acquisition of phonotactic regularities. *Journal of psycholinguistic research 38*(2), 129–150.

LaRiviere, C., H. Winitz, J. Reeds, and E. Herriman (1974). The conceptual reality of selected distinctive features. *Journal of Speech and Hearing Research 17*(1), 122–133.

LaRiviere, C., H. Winitz, J. Reeds, and E. Herriman (1977). Erratum: The conceptual reality of selected distinctive features. *Journal of Speech and Hearing Research 20*(4), 817.

Lindsay, S. and M. G. Gaskell (2010). A complementary systems account of word learning in L1 and L2. *Language Learning 60*(Suppl. 2), 45–63.

Love, B. C., D. L. Medin, and T. M. Gureckis (2004). SUSTAIN: a network model of category learning. *Psychological Review 111*(2), 309–332.

Mackie, S. and J. Mielke (2011). Feature economy in natural, random, and synthetic inventories. In G. N. Clements and R. Ridouane (Eds.), *Where do phonological features come from? Cognitive, physical, and developmental bases of distinctive speech categories*, pp. 43–63. John Benjamins.

Mattys, S. L. and P. W. Jusczyk (2001). Phonotactic cues for segmentation of fleunt speech by infants. *Cognition 78*(2), 91–121.

McCarthy, J. J. (1988). Feature Geometry and dependency: a review. *Phonetica 45*, 84–108.

Mielke, J. (2004). *The emergence of distinctive features.* Ph. D. thesis, Ohio State University.

Moreton, E. (2008). Analytic bias and phonological typology. *Phonology 25*(1), 83–127.

Moreton, E. (2010). Underphonologization and modularity bias. In S. Parker (Ed.), *Phonological argumentation: essays on evidence and motivation*, Chapter 3, pp. 79–101. London: Equinox.

Moreton, E. (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language 67*(1), 165–183.

Moreton, E., G. Feng, and J. L. Smith (2008). Syllabification, sonority, and perception: new evidence from a language game. In R. L. Edwards, P. J. Midtlyng, C. L. Sprague, and K. G. Stensrud (Eds.), *Papers from the 41st Regional Meeting of the Chicago Linguistic Society*, pp. 341–355.

Nevins, A. (2010). Two case studies in phonological universals: a view from artificial grammars. *Biolinguistics 4*(2), 218–233.

Nosofsky, R. M., T. J. Palmeri, and S. C. McKinley (1994). Rule-plus-exception model of classification learning. *Psychological Review 101*(1), 53–79.

Ohala, J. J. (1992). What's cognitive, what's not, in sound change. In G. Kellerman and M. D. Morissey (Eds.), *Diachrony within synchrony: language history and cognition*, pp. 309–355. Frankfurt am Main: Peter Lang.

Ohala, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: problems and perspectives*, pp. 237–278. Harlow: Longman.

Ohala, J. J. (1997). Phonetics in phonology. In *Proceedings of the 4th Seoul International Conference on Linguistics (SICOL), 11–15 August*, pp. 45–50.

Padgett, J. (2002a). Feature classes in phonology. *Language 78*(1), 81–110.

Padgett, J. (2002b). The unabridged "feature classes in phonology". MS, Department of Linguistics, University of California, Santa Cruz.

Pater, J. and E. Moreton (in press). Structurally biased phonology: complexity in learning and typology. *Journal of the English and Foreign Languages University, Hyderabad*.

Pater, J. and A.-M. Tessier (2006). L1 phonotactic knowledge and the L2 acquisition of alternations. In R. Slabakova, S. Motrul, P. Prévost, and L. White (Eds.), *Inquiries in linguistic development: in honor of Lydia White*, pp. 115–131. Benjamins.

Peperkamp, S., K. Skoruppa, and E. Dupoux (2006). The role of phonetic naturalness in phonological rule acquisition. In D. Bamman, T. Magnitskaia, and C. Zoller (Eds.), *Papers from the 30th Boston University Conference on Language Development (BUCLD 30)*, pp. 464–475.

Pertz, D. L. and T. G. Bever (1975). Sensitivity to phonological universals in children and adults. *Language 51*, 149–162.

Picard, M. (2003). On the emergence and resolution of hiatus. *Folia Linguistica Historica 24*(1–2), 47–57.

Pons, F. and J. M. Toro (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition 116*, 361–367.

Pulleyblank, D. (1988). Underspecification, the feature hierarchy, and Tiv vowels. *Phonology 5*(2), 299–326.

Pycha, A., P. Nowak, E. Shin, and R. Shosted (2003). Phonological rule-learing and its implications for a theory of vowel harmony. In M. Tsujimura and G. Garding (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics (WCCFL 22)*, pp. 101–114.

Rose, S. and R. Walker (2004). A typology of consonant agreement as correspondence. *Language 80*(3), 475–531.

Saffran, J. R. and E. D. Thiessen (2003). Pattern induction by infant language learners. *Developmental Psychology 39*(3), 484–494.

Schane, S. A., B. Tranel, and H. Lane (1974). On the psychological reality of a natural rule of syllable structure. *Cognition 3*(4), 351–358.

Simons-Wiedenmaier, N. S., M. Weber, C. F. Plappert, P. K. D. Pilz, and S. Schmid (2006). Synaptic depression and short-term habituation are located in the sensory part of the mammalian startle pathway. *BMC Neuroscience 7*(38).

Skoruppa, K., A. Lambrechts, and S. Peperkamp (2009). The role of phonetic distance in the acquisition of phonological alternations. In S. Lima, K. Mullin, and B. Smith (Eds.), *Proceedings of the 39th Meeting of the North-East Linguistics Society (NELS)*, Amherst, Massachusetts. Graduate Linguistics Students Association.

Skoruppa, K. and S. Peperkamp (2011). Adaptation to novel accents: feature-based learning of context-sensitive phonological regularities. *Cognitive Science 35*, 348–366.

St. Clair, M. C. and P. Monaghan (2008). Language abstraction: consolidation of language structure during sleep. In B. C. Love, K. McRae, and V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society*, Austin, Texas, pp. 727–732. Cognitive Science Society.

Stampe, D. (1973). *A dissertation on Natural Phonology*. Ph. D. thesis, University of Chicago.

Steriade, D. (1997). Phonetics in phonology: the case of laryngeal neutralization. MS, Department of Linguistics, University of California, Los Angeles.

Steriade, D. (2001). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. MS, Department of Linguistics, University of California, Los Angeles.

Taylor, C. F. and G. Houghton (2005). Learning artificial phonotactic constraints: time course, durability, and relationship to natural constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition 31*, 1398–1416.

Toro, J. M., M. Nespor, J. Mehler, and L. L. Bonatti (2008). Finding words and rules in a speech stream. *Psychological Science 19*(2), 137–144.

Toro, J. M., M. Shukla, M. Nespor, and A. D. Endress (2008). The quest for generalizations over consonants: asymmetries between consonants and vowels are not the by-product of acoustic differences. *Perception and Psychophysics 70*(8), 1515–1525.

van Kampen, A., G. Parmaksiz, R. van de Vijver, and B. Höhle (2008). Metrical and statistical cues for word segmentation: vowel harmony and word stress as cues to word boundaries by 6- and 9-month-old Turkish learners. In A. Gavarroó and M. J. Freitas (Eds.), *Language Acquisition and Development*, pp. 313–324. Cambridge Scholars Publishing.

Wagner, U., S. Gais, H. Haider, R. Verleger, and J. Born (2004). Sleep inspires insight. *Nature 427*, 352–355.

Weinreich, U., W. Labov, and M. Herzog (1968). Empirical foundations for a theory of language change. In W. P. Lehmann and Y. Malkiel (Eds.), *Directions for historical*

linguistics: a symposium, pp. 97–195. Austin: University of Texas Press.
Whalen, D. H. and A. G. Levitt (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics 23*, 349–366.
Wilson, C. (2001). Consonant cluster neutralisation and targeted constraints. *Phonology 18*, 147–197.
Wilson, C. (2003). Experimental investigation of phonological naturalness. In G. Garding and M. Tsujimura (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics (WCCFL 22)*, Somerville, pp. 533–546. Cascadilla Press.
Wilson, C. (2006). Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science 30*(5), 945–982.
Yu, A. C. L. (2012). On measuring phonetic precursor robustness: a response to Moreton. *Phonology 28*(3), 491–518.
Zaba, A. M. (2008). *Relative frequency of patterns and learnability: the case of phonological harmony*. Ph. D. thesis, University of Utah.
Zhang, J. and Y. Lai (2010). Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology 27*(1), 153–201.
Zhang, J., Y. Lai, and C. Sailor (2011). Modeling Taiwanese speakers' knowledge of tone sandhi in reduplication. *Lingua 121*(2), 181–206.
Zhang, J., Y. Lai, and C. Turnbull-Sailor (2006). Wug-testing the "tone circle" in Taiwanese. In D. Baumer, D. Montero, and M. Scanlon (Eds.), *Proceedings of the 25th meeting of the West Coast Conference on Formal Linguistics (WCCFL)*, Somerville, Mass., pp. 453–461. Cascadilla Proceedings Project.
Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: corpus and survey evidence from Tagalog infixation. *Language 83*(2), 277–316.