

# NIH Public Access

**Author Manuscript** 

Cogn Sci. Author manuscript; available in PMC 2013 May 01.

#### Published in final edited form as:

Cogn Sci. 2012 May ; 36(4): 740-756. doi:10.1111/j.1551-6709.2011.01227.x.

# Testing the limits of long-distance learning: Learning beyond a three-segment window

#### Sara Finley

Department of Brain and Cognitive Sciences, University of Rochester, Meliora Hall, Rochester, NY 14627, sfinley@bcs.rochester.edu

# Abstract

Traditional flat-structured bigram and trigram models of phonotactics are useful because they capture a large number of facts about phonological processes. Additionally, these models predict that local interactions should be easier to learn than long-distance ones since long-distance dependencies are difficult to capture with these models. Long-distance phonotactic patterns have been observed by linguists in many languages, who have proposed different kinds of models, including feature-based bigram and trigram models, as well as precedence models. Contrary to flat-structured bigram and trigram models, these alternatives capture unbounded dependencies because at an abstract level of representation, the relevant elements are locally dependent, even if they are not adjacent at the observable level. Using an artificial grammar learning paradigm, we provide additional support for these alternative models of phonotactics. Participants in two experiments were exposed to a long-distance consonant harmony pattern in which the first consonant of a five-syllable word was [s] or  $[\int]$  ('sh') and triggered a suffix that was either [-su] or  $[-\int u]$  depending on the sibilant quality of this first consonant. Participants learned this pattern, despite the large distance between the trigger and the target, suggesting that when participants learn long-distance phonological patterns, that pattern is learned without specific reference to distance.

### Keywords

phonology; phonotactics; statistical learning; n-gram models; consonant harmony

# 1. Introduction

An important aspect of human cognition is how patterns and associations are learned at a distance. In language, the vast majority of patterns occur between adjacent elements (Chomsky, 1981; Culicover & Wilkins, 1984). However, there are a large number of apparent exceptions to this generalization. These long-distance patterns pose important challenges for both theoretical and computational models of human cognition (Gomez, 2002). While the majority of computational and experimental work concerning long-distance phenomena has focused on the sentence level (Misyak & Christiansen, 2007; Misyak, Christiansen, & Tomblin, 2009; Onnis, Christiansen, Chater, & Gomez, 2003; Onnis, Monaghan, Christiansen, & Chater, 2004), the same issues apply at the word level. Specifically, there are important challenges for explaining long-distance patterns in phonotactics, the rules that govern possible sound sequences in a language.

Phonotactic patterns in language have three levels of locality. The most local level is adjacency; patterns apply to sounds that are adjacent. In nasal place assimilation, a coronal

Finley

nasal takes the place of the following, adjacent stop consonant (e.g., the phrase 'run pool' would be pronounced as 'rum pool' after nasal place assimilation, but 'run open' remains unchanged because the [n] and the [p] are non-adjacent). The second level of locality, referred to in this paper as first-order non-locality, is when the segments involved in the pattern need not be adjacent, but they are local on the level of vowel or consonant. For example, in vowel harmony, vowels must agree in feature values with the nearest vowels to the left or right, but these vowels need not be adjacent, as consonants can intervene. The most extreme cases of non-locality in phonotactics referred to in the present paper as second-order non-locality, (often referred to as 'unbouded'1), occur when a phonotactic pattern can apply across distances that allow for intervening consonants and vowels.

The differences between first-order and second-order phonotactics are most clearly demonstrated in consonant harmony. Consonant harmony is a phonotactic pattern in which consonants must agree in some phonological feature, such as place of articulation. One of the most common varieties of consonant harmony involves sibilant fricatives [s] and [ $\int$ ] ('sh'). In sibilant harmony, [s] and [ $\int$ ] alternate such that if a word contains the sound [ $\int$ ], all preceding (or following) [s] sounds become [ $\int$ ]. For example, Navajo shows alternations in prefixes between [si-] and [ $\int$ i-] (Hansson, 2001; Heinz, 2010; McDonough, 1990; Rose & Walker, 2004; Sapir & Hojier, 1967). The prefix [si-] appears if the stem contains the alveolar fricatives /s, z, ts, ts', dz/, as in [site:z] 'my car', while [ $\int$ i-] appears if the stem contains the postalveolar fricatives / $\int$ , t $\int$  ('ch'), z, dz/, as in [ $\int$ it $\int$ ii:h] 'my nose'.

In first-order long-distance consonant harmony, a vowel can readily intervene between two sibilant consonants (they need not be directly adjacent). In the Navajo example above, [s] and [ $\int$ ] alternate in the perfective prefix, but the vowel [i] remains constant ([si-] versus [ $\int$ i-]). In second-order long-distance consonant harmony, irrelevant consonants are ignored. For example, the form [ $\int$ ite:3] 'we are lying' involves agreement between the first and last consonants, skipping the middle consonant. The distances between the two relevant consonants can be quite large. Hannsson cites several examples from Chumash, which has a similar sibilant harmony system to Navajo, in which agreement occurs across several sound segments (e.g., [ha $\int$ xintilawa $\int$ ] 'his former Indian name', in which nine segments (four of which are vowels) intervene between the two sibilant consonants.

Traditional linguistic models of first and second order non-locality use hierarchical, tierbased structures to explain the differences between local and non-local processes in phonology (Archangeli & Pulleyblank, 1994; Booij, 1984; Clements, 1976, 1977; Goldsmith, 1975). Tier-based models of phonotactics are based on the assumption that linguistic structures are hierarchically organized2. Tier-based models maintain a theoretical notion of adjacency because different segments of particular feature values are placed onto different representational dimensions (or 'tiers'). In Figure 1, the only sibilant consonants in  $[\int ite:3]$  are  $[\int i]$  and [3]. Because sibilants are given their own tier, the agreement of features can be represented locally, despite the fact that the stop consonant [t] intervenes on the surface.

The hierarchically structured tiers-based model posed by theoretical linguists significantly contrasts with traditional computational models of phonotactics that employ a flat-structured representation for segments. The most common computational method for capturing linguistic data uses n-gram models (Jurafsky & Martin, 2008). The basic premise behind an

<sup>&</sup>lt;sup>1</sup>It is important to note that the idea of 'unbounded' in linguistics is a theoretical notion. In theory, a word or sentence may be infinitely long, but in practice, cognitive constraints on attention and memory limit the size of words.

infinitely long, but in practice, cognitive constraints on attention and memory limit the size of words. <sup>2</sup>For example, an autosegmental, tier-based analysis of the root and pattern morphology found in Arabic words would break up the word [samam] with a root [sm] and the inflection [a], each on a different tier.

n-gram model of phonotactics is that words are divided into chunks of length n (e.g., length of 2 for bigrams, length of 3 for trigrams, etc.). For example, the word [gredId] 'graded 'has four trigrams: [gre], [red], [edI] and [dId]. Statistics calculated over these strings determines the restrictions of sound sequences in words. For example, the restriction against words in English beginning with *pt* is captured by the fact that there are no word-initial *pt* bigrams. One of the major appeals of n-gram models is that n-gram models do not require abstract and hierarchical information to model phonotactic patterns. Because the majority of phonotactic patterns occur within a set of three adjacent segments, trigrams are sufficient to capture most phonotactic patterns (Hayes & Wilson, 2008; Heinz, 2007). In addition, n-gram models, while not intended to be a true model for human cognition, tend to outperform models based on linguistic theory (Chelba & Jelinek, 2000; Jurafsky & Martin, 2008).

It is unclear whether n-gram models are capable of accounting for long-distance patterns without reference to structural elements (Gomez, 2002; Hayes & Wilson, 2008; Heinz, 2007; Onnis, et al., 2003). First, modeling long-distance patterns with large n-grams reduces the computational simplicity of the model, and creates a data sparsity problem. As n increases, the statistics become less useful and require more powerful computations (Jurafsky & Martin, 2008). Second, there is a question of how n should be chosen. One could store all possible n-grams (up to the length of the longest word in the language), or limit the algorithm to a particular value of n at the risk of missing patterns active at longer distances. This problem of a principle for limiting traditional n-gram statistics leads to a hypothesis that if n-gram statistics bear psychological reality, they are not the sole method for learning and representing data. Further, it leads one to expect that n-gram models must be buttressed by some other mechanism to account for long-distance patterns, as well as the principled nature of phonotactics (e.g., the fact that the majority of phonotactic restrictions are based on phonetic principles).

One possibility is to incorporate a tier-based structure (Archangeli & Pulleyblank, 1994; Booij, 1984; Clements, 1976, 1977; Goldsmith, 1975) into an n-gram model of phonotactics (Hayes & Wilson, 2008). An n-gram model without tiers can describe a first-order longdistance dependency, because the number of segments intervening between non-adjacent elements is restricted ion most cases. However, n-gram models without tiers cannot accurately account for second-order long-distance dependencies because some consonants (or vowels in the case of vowel harmony) will also have to be erased before an n-gram model can be effective. The addition of tiers into an n-gram model would allow unbounded agreement between sibilants to be modeled with bigrams on a sibilant tier, as well as other principles of phonotactics3.

Another alternative to traditional n-gram models is to model precedence relationships (Hayes & Wilson, 2008; Heinz, 2010; Heinz & Rogers, 2010). In precedence models, restrictions on sound sequences are based on frequencies of pairs of sounds that need not be adjacent, but must be in a precedence relation (e.g., *ts* but not *st* in *trucks*). Precedence relations are stated in terms of two segments, but these segments need not be adjacent. For example, the consonants in the form [sokibosu] have the following precedence relations: [s-k], [s-b], [s-s], [k-b], [k-s], [b-s]4. By calculating statistics over the precedence relationships, the model can find the principled representations. For example, over a corpus of items in a sibilant harmony language, the model will find precedence relationships [s-s] and  $[\int - \int]$ , but not [s- $\int$  and  $[\int -s]$ , suggesting a sibilant harmony constraint. Precedence relations do not

<sup>&</sup>lt;sup>3</sup>Note that a richer set of structural principles would have to be implemented in order to account for transparent vowels in vowel harmony, particularly in Hungarian, which allows for multiple transparent vowels (Hayes & Londe, 2006).

<sup>&</sup>lt;sup>4</sup>Heinz (2010) distinguishes between adjacent and non-adjacent precedence relations with ellipses (e.g., [pz] denotes an adjacent relationship (e.g., 'pzat'), while [p...z] denotes a non-adjacent relationship (e.g., 'pizat'). Without this distinction the precedence model cannot distinguish between first-order phonotactics and second-order.

encode distance, meaning that segments need not be adjacent or separated by featural representations in order to be encoded in the model. This allows the model to represent unbounded phonotactic processes. It is important to note, however, that these alternatives are designed to capture long-distance patterns, not adjacent patterns. As Heinz (2010) notes, the precedence model is designed to supplement a strictly 2-local learner, rather than supplant it.

The present study tests the degree to which learners encode and represent novel longdistance phonotactic patterns. Using an artificial grammar learning paradigm, it is possible to directly study the mechanisms that underlie biases in adult human learning. The test phase of an artificial grammar learning experiment can be used to determine what learners infer from the material that they are exposed to. By testing participants on material that they have not heard before, it is possible to discern the level of generality with which particular patterns are learned (Finley & Badecker, 2009a; Wilson, 2006).

Previous research has shown that adults can learn vowel harmony (Finley & Badecker, 2008, 2009a, 2009b, in press; Pycha, Nowak, Shin, & Shosted, 2003) and consonant harmony patterns (Finley, 2011; Wilson, 2003) with a relatively short amount of training. Thus, the first-order long-distance dependency of 'skipping' vowels or consonants is not a problem for learners in an artificial grammar learning setting. Finley (in submission) showed that second-order non-adjacent dependencies in vowel harmony are more difficult to learn than local counterparts. However, the representations required for long-distance dependencies in vowel harmony tend to be structurally more complex than those required for consonant harmony. In Finley (in submission), learners were exposed to a vowel harmony pattern that, in addition to the regular first-order non-adjacency, contained an intervening vowel that was inert to harmony; the harmonic feature passed through this vowel, and the feature value of the suffix vowel was determined non-locally. Participants were only able to learn the non-local pattern after a significant increase in training and tokens of test items. This suggests that long-distance patterns can be difficult for learners when such patterns require complex representations. Consonant harmony, unlike vowel harmony, does not require the same complex representations. The reason for this difference lies in the phonetic features of intervening consonants. Unlike the features of intervening vowels in vowel harmony, the features of intervening consonants in consonant harmony are irrelevant to the specific features involved in harmony, and can be separated onto different structural tiers (see Figure 1). In addition, consonants in harmony tend to be underspecified, such that the phonetic implementation can be gleaned via interpolation (Keating, 1988). Because there are fewer features distinguishing vowels from other vowels, intervening vowels in vowel harmony require more complex representations to sort out which vowels participate and which do not.

Finley (2011) compared learning of a first-order long-distance harmony pattern, in which the two relevant consonants were separated only by a vowel (e.g., besosu), with learning a second-order long-distance harmony pattern, in which vowels and consonants separated the two relevant consonants (e.g., sobesu). If a language requires second-order dependencies, it will also require first-order dependencies, but there are languages that require first-order dependencies without requiring second-order dependencies (Hansson, 2001). Learners in Finley (2011) followed the same pattern in their inferences in the artificial consonant harmony patterns; when exposed to a second-order pattern (e.g., sobesu), learners generalized to a first-order pattern (e.g., besosu), but when exposed to a first-order pattern (e.g., sobesu). This suggests that learners prefer a first-order pattern if it can explain all of the data, but will infer a general, second-order pattern otherwise.

While Finley (2011) demonstrated long-distance learning of consonant harmony, all of the items in those experiments could be analyzed within a three-segment consonant window (e.g. sobesu as sbs). It is unclear what occurs when the distance between the relevant consonants increases beyond a three-segment window. If learners make use of non-hierarchical, flat representations of the input, one should expect that the learning problem becomes more difficult as the distance between the relevant segments increases. While it is generally understood that n-grams are a rough approximation for human language processing, it is important to understand the ways in which the implications that such models make regarding the nature of the learning problem are appropriate for modeling human inferences in learning.

In the present study, we extend the artificial grammar learning paradigms used in previous harmony learning experiments. Experiment 1 tests whether it is possible for adults to learn long-distance consonant patterns that involve multiple consonant interveners. Experiment 2 tests whether learners exposed to a consonant harmony pattern that takes place within a three-segment window will be able to extend the pattern to forms that have additional intervening consonants and are therefore outside a three-segment window. We show that adults can learn phonotactic patterns across multiple intervening segments, and do not appear to encode exact number of intervening segments in pattern learning.

## 2. EXPERIMENT 1

Experiment 1 tests whether adult English speakers are able to learn a consonant harmony pattern that takes place across multiple intervening consonants. In this experiment, learners were auditorily exposed to a consonant harmony pattern in which a suffix alternated between [-su] and  $[-\int u]$  depending on whether the first consonant of the word was [s] or  $[\int]$ . The form of the suffix was determined by the first vowel of a tri-syllabic stem. This ensured that there were two syllables intervening between the triggering consonant and the suffix (e.g., [sobigosu]). Following exposure, participants were tested on familiar and unfamiliar suffixed forms, assessing the level of learning of the consonant harmony patterns. If learners are able to learn long-distance consonant harmony pattern. Further, if learners form a general, second-order long-distance harmony pattern, learners should generalize the harmony pattern to cases that have fewer intervening segments.

#### 2.1. Method

**2.1.1. Participants**—All participants were adult native English speakers with no knowledge of a consonant harmony language, and had not previously participated in a consonant harmony learning experiment. Fifty-four University of Rochester undergraduate students and affiliates and were paid \$10 for participation.

**2.1.2. Design**—The experiment consisted of a training phase followed immediately by a forced-choice test. All phases of the experiment were presented using PsyScopeX (Cohen, MacWhinney, Flatt, & Provost, 1993). All stimuli were presented auditorily.

Participants in the Critical condition were exposed to a second-order long-distance consonant harmony language characterized by an alternation between two suffix allomorphs [-su] and  $[-\int u]$ ; [-su]. The allomorphs were conditioned on the sibilant consonant in a tri-syllabic stem form. 12 stems were of the form sVCVCV, and 12 stems were of the form  $\int V$  CVCV; [-su] surfaced when the initial consonant was [s] and  $[-\int u]$  surfaced when the initial consonant was [s] and [- $\int u$ ] surfaced when the initial consonant was [s]. All items were presented in pairs: stem followed by stem+suffix (e.g., [sokubi~sokubisu], with a 500ms pause between the stem and the stem+suffix form. The first consonant in each stem was either [s] or [ $\int$ ], and the additional consonants were

stops [p, t, k, b, d, g]. Vowels were drawn from the set [a, i, e, o, u]. There were 24 stemsuffix pairs repeated in a random order 5 times each. Examples of stimuli can be found in Table 1.

A Control condition was created to assess for biases in the stimuli and/or prior to the experiment. In the Control condition, learners were exposed to the same sVCVCV and  $\int VCVCV$  stems as those heard in the Critical condition. However, participants did not hear the suffixed form, which gave them no access to a harmony pattern. Providing participants with some training makes it possible to give the same instructions (training and test), as the Critical condition (as opposed to a 'no training' Control condition5), and allows us to assess the role of the stem in influencing learners' responses. It is possible that the sVCVCV and  $\int VCVCV$  nature of the stems could serve to bias learners without actual training on these items. Participants in the Control condition were given identical test items to those in the Critical condition, making it possible to directly compare the Critical and the Control conditions.

Following training, participants were given a two-alternative forced-choice test in which participants chose between two stem+suffix forms. Both alternatives contained the same stem, but the suffix item was either [-su] and  $[-\int u]$  (with a 500ms pause between each stem +suffix option), meaning that only one item would be harmonic (e.g., harmonic [sokebusu] was pit against disharmonic [sokebo∫u]). Test items included Old Items6 (in which the stems were drawn from the training set), New Items (in which the stems were not drawn from the training set, but were of the same sVCVCV and  $\int VCVCV$  form seen at training), and Trigram Items. Trigram Items tested learners' ability to generalize the consonant harmony pattern to a different number of intervening segments. The sibilant in Trigram Items always appeared as the second consonant in the stem. Trigram test items were of the form \*7CV∫CVsu vs. CV∫VCV∫u. Given the salience of the fact that all items in the training set began with a sibilant consonant, we expect some degraded performance for Trigram Items. The full stimuli list for Experiment 1 can be found in Appendix 1.

If participants in the Critical condition learned the general harmony pattern they should select the correct allomorph at a rate significantly greater than participants in the Control condition.

**2.1.3. Materials**—All stimuli were recorded by an adult female native English speaker in a sound-attenuated booth. While the volunteer was aware that the stimuli would be used in an artificial grammar learning experiment, she was blind to the design or purpose of the study. The speaker produced all vowel sounds without reduction, and stress was placed on the first syllable; stress was consistent across stem and suffixed items.

#### 2.3 Procedure

All participants were given written and verbal instructions. Participants were told that they would be listening to words from a language they never heard before, and that their task was to listen to the way the novel language sounded, but that they need not memorize the forms. The training was followed by a forced-choice test with 36 pairs of suffixed items, one item in the pair harmonic and the other item disharmonic (e.g., [\*sibopi $\int u$  vs.  $\int ibopi \int u$ ]).

<sup>&</sup>lt;sup>5</sup>Previous research (e.g., Finley and Badecker (2009)) found no differences between 'no-training' Control conditions and stem-only Control conditions.

<sup>&</sup>lt;sup>6</sup>All items were technically 'new' for the Control condition, as these participants did not hear any suffixed items. While participants in the Critical condition could respond to Old Items based on familiarity, both harmonic and disharmonic options were equally unfamiliar to participants in the Control condition. <sup>7</sup>The '\*' indicates the disharmonic item.

given a debriefing statement upon completion of the experiment (which took approximately 15 minutes).

#### 2.2. Results

Proportions of harmonic responses were recorded for each subject in the Critical and the Control conditions, and can be found in Figure 2. Participants in the Control condition were compared to participants in the Critical condition via a mixed design ANOVA with alpha set at .05. The between-subjects factor was Training, with two levels: Critical and Control. Test Items (Old Items, New Items, Trigram Items) was a within-subjects factor nested under the between-subjects factor Training. All conditions involved between-item comparisons.

There was a significant effect of Training (F (1,52) =17.21, p < 0.001), in that participants in the Critical condition were more likely to choose the harmonic option than participants in the Control condition (mean = 0.69 vs. 0.50, *CI*±0.093). There was a significant interaction (*F*(2, 104) = 2.98, p=0.054). This interaction is due to the fact that in there was a significant difference between New and Trigram Items for the Critical condition (t(53) = 2.85, p = 0.006), but not the Control condition (t(53) = 1.11, p = 0.28). There was a marginal effect of Test Item (*F*(2, 104) = 4.62, p = 0.082), which was carried by the fact that there were significant differences between New and Trigram Items (F(1, 52) = 8.72, p = 0.005) and Old and Trigram Items (*F*(1, 52) = 4.04, p = 0.050).

To directly test whether participants generalized to the Trigram Items, we performed a t-test between Critical and Control conditions for the Trigram Items. There was a significant difference between the Critical condition and the Control condition (t(52) = 2.36, p = 0.02), further suggesting that participants extended the harmony pattern to the Trigram Items. There was, however, a significant difference between New and Trigram Items for the Critical condition (t(26)=3.02, p=0.006). Given that the initial position was relevant to the long-distance items, but not to the Trigram Items, it is likely that some participants may have inferred that the harmony pattern only applied when the initial consonant was a sibilant. However, there are two reasons to believe that a minority of participants made this inference. First, there was significant generalization to trigram Items. Second, of the 12 participants who scored at or below chance on Trigram items, only 3 scored above 50% on Old Items, suggesting that these participants failed to learn the harmony pattern. None of these participants showed any preference for suffixes containing  $/\int/$  compared to suffixes containing /s/(t<1).

#### 2.3. Discussion

Participants in Experiment 1 were able to learn the consonant harmony pattern at a distance beyond the three-segment window, supporting approaches to long-distance patterns that do not specifically encode distance. Learners were able to extend the pattern to the more local Trigram test items. While participants were less robust at generalization to the Trigram Items, this should be expected given that all of the items in the training set began with a sibilant, but none of the Trigram Items did. What is important is that there was a significant difference between the Critical and Control conditions for the Trigram Items. This suggests that learners did not specifically encode the number of intervening segments between the first and the last consonants.

It is possible that learners only form general representations when a trigram representation is insufficient. If learners use the three-segment window to learn and encode novel long-distance patterns, a learner who was exposed only to input that could be captured by a three-

segment window would be unable to generalize the pattern to items that involve a greater number of intervening segments. Further, because the initial consonant was highly relevant in Experiment 1, it is possible that learners were able to encode the long-distance pattern only because of the position of the relevant consonants. These possibilities are tested in Experiment 2; learners were exposed to the long-distance sibilant harmony pattern within the three-segment window, and the initial consonant is not relevant for the harmony pattern. Following exposure, learners were given a set of test items (parallel to Experiment 1) that included more than two segments intervening between the two sibilant consonants. If learners extend the harmony pattern to these 'Long-Distance' test items, they have learned a general harmony pattern that does not encode distance.

# 3. EXPERIMENT 2

#### 3.1. Method

**3.1.1. Participants**—All participants were adult native English speakers with no knowledge of a consonant harmony language, and had not previously participated in a consonant-harmony learning experiment. Forty-eight University of Rochester undergraduate students and affiliates and were paid \$10 for participation. There were 24 participants in each condition.

**3.1.2. Design**—The design of Experiment 2 was identical to Experiment 1, except that learners were exposed to a long-distance harmony pattern in training that fit within the three-segment consonant window, and were tested on generalization to a long-distance pattern that extended the number of intervening segments between the two sibilant consonants. Examples of training and test stimuli can be found in Tables 1 and 2.

**3.1.3. Materials**—The materials were recorded in the same manner as Experiment 1. The full stimuli list for Experiment 2 can be found in Appendix 2.

**3.1.4. Procedure**—The procedure was identical to Experiment 1.

#### 3.2. Results

Proportions of harmonic responses were recorded for each subject in the critical and the control conditions, and can be found in Figure 3. Participants in the Control condition were compared to participants in the Critical condition via a mixed design ANOVA with alpha set at .05. The between-subjects factor was Training: Critical and Control. Test Items (Old Stems, New Stems, Long Distance) was a within-subjects factor nested under the between-subjects factor Training. All conditions involved between-item comparisons.

There was a significant effect of Training (R(1,46) = 22.56, p < 0.0001), in that participants in the Critical condition were more likely to choose the harmonic option than participants in the Control condition (mean = 0.69 vs. 0.45,  $CI \pm 0.10$ ). There was no effect of Test Item (F < 1) and no interaction (R(2, 92) = 1.98, p = 0.14).

To directly test whether participants generalized to long-distance patterns, we performed a ttest between Critical and Control conditions for the Long Distance items. There was a significant difference (t(46)=2.98, p=0.005), suggesting that learners generalized from the trigram training to a long-distance pattern. Further, there was no difference between New and Long Distance test items (t<1).

Because distance and position are confounded in the present study (i.e., the long-distance items always contained a sibilant as the first consonant in the word), we performed several

Page 8

cross-experiment comparisons to ensure that the role of distance and position did not differ. We compared Experiment 1 and Experiment 2 for responses to New (t<1) and New Distance (Trigram for Experiment 1 and Long-Distance for Experiment 2) (t<1) and found no differences. We also compared across Experiments 1 and 2 for the same items. We compared Trigram Items from Experiment 1 with New Items from Experiment 2 and found no differences (t<1), and New Items from Experiment 1 with Long-Distance Items from Experiment 2 and found no differences (t<1). This suggests that position did not directly affect the level of generalization to novel distances.

Of the 24 participants in the Critical condition of Experiment 2, only 5 participants were below chance for the Long-Distance items, and 6 were at chance. Of these 11 participants, only 3 were above 60% for New Items, suggesting that these participants simply failed to learn the harmony pattern. None of these participants showed any preference for suffixes containing  $/\int/$  compared to suffixes containing /s/(t<1).

#### 3.3. Discussion

Participants in Experiment 2 learned the long-distance consonant harmony pattern, and extended that harmony pattern to cases requiring an even greater distance between the two sibilant consonants. Increasing the number of intervening consonants did not affect learners' acceptance of harmonic strings at a distance. These results provide further support for a view that learners encode long-distance consonant harmony patterns without reference to distance, in accordance with a theoretical notion of 'unboundedness'.

# 4. GENERAL DISCUSSION

This study presented two experiments exhibiting adult humans' ability to learn long-distance phonotactics at relatively great distances. When exposed to a consonant harmony pattern in which the first sibilant consonant of a word determined the final sibilant consonant of the word, participants could learn such patterns even when two syllables intervened between the two sibilant consonants. Further, learners extended the consonant harmony pattern to forms of varying numbers of intervening elements (fewer elements in Experiment 1 and more elements in Experiment 2).

These results provide important information about how phonotactic restrictions are learned and represented. Learners appear to be able to learn long-distance patterns with relative ease. If learners gave privileged status to interactions that can take place at the bigram or trigram level, one might expect that learners would display degraded performance as the number of grams required to represent the pattern increases. However, learners appear to be resilient to increases in the number of interveners. Participants in Experiment 1 and Experiment 2 showed relatively equal levels of learning with mean harmonic responses of 68% for Experiment 1 and 69% for Experiment 2. If learners had access only to trigram representations, not only would participants have failed to learn the long-distance pattern in Experiment 1, they should have showed degraded performance as distance increased, but this did not occur. One interesting pattern in the results is that there seemed to be less robust generalization from long-distance to trigram items in Experiment 1 than the reverse generalization in Experiment 2. This is the opposite pattern that would occur if learners only used distance information in encoding phonological patterns. It is possible that learners made use of positional information.

The results demonstrate that bigrams and trigrams do not appear to be privileged in learning. This does not mean that learners are never sensitive to bigram and trigram information when they are relevant. Rather, it appears that learners are sensitive to a wider set of statistics. Further, while the results are consistent with an unbounded phonotactic pattern, they are also

consistent with a model that encodes up to four consonant n-grams. While it might be possible to show learning at greater distances, it is in general impossible to experimentally demonstrate the learning of truly unbounded patterns. This paper supports the notion that learners are not strictly bound to distance, but can generalize beyond the distance that they are exposed to. Note that this notion of unboundedness applies only to second-order long distance patterns. As Finley (2011) demonstrated, learners of a consonant harmony pattern that applied without any consonants intervening between the relevant consonants did not generalize to a second-order long-distance pattern with consonants and vowels intervening.

The results of the present study raise the question of how best to measure models of phonotactic learning. While it is often assumed that the model that can capture the most amount of data with the simplest representations is superior, the results of the experiments presented in this paper suggest that if models are to be used as a tool for understanding the workings of the mind, computational models of cognition must take into account how human learners actually behave. Traditional n-gram models may capture a large proportion of human performance with relatively fewer computational constraints, but such models do not capture the intuitions of human learners or their behavior in artificial language learning experiments like this one. On the other hand, hierarchically structured n-gram and precedence models have the power to account for long-distance phonotactic patterns without the sparsity issues encountered in traditional n-gram models. While the present experiment is consistent with the predictions of these models, it does not distinguish between them.

This suggests that goodness of fit and generalization as measures of a model's success do not provide the most complete measures of a model's success. In order to evaluate computational models of cognition, experimental and computational approaches must work together. Incorporating experimental results into the implementation of computational models will invariably lead to a better understanding of how the mind is constrained, and how such constraints might be implemented.

While the present study addressed only the consequences for incorporating long-distance dependencies into n-gram models of phonotactics, there are important consequences for models of other areas of language. Because long-distance dependencies can be found at the word level, the morpheme level and the sentence level, models of a wide range of linguistic processes may be affected by their ability to account for processing long-distance processes. For example, Newport and Aslin (2004) found that adult learners are able to segment words from speech that contained high transitional probabilities for non-adjacent segments (but not from non-adjacent syllables). This biased ability to segment speech at a distance may have important consequences for n-gram models of word segmentation (Cairns, et al., 1997). Future research will work to establish the accuracy of n-gram models in word segmentation, as well as other linguistic areas. For example, word-level trigrams have been used as a basis for learning grammatical categories (St. Clair, Monaghan, & Ramscar, 2009). It would be interesting to understand the extent to which word-level trigram models are affected by long-distance patterns at the sentence level.

# **5. CONCLUSIONS**

The present study demonstrated that human adults are able to learn a novel harmony pattern at a distance. Participants were able to learn the harmony pattern even when the pattern fell outside a three-segment, trigram window. These results suggest that learners are not sensitive to distance when learning long-distance patterns, and that consonant harmony patterns are consistent with the notions developed in theoretical linguistics that long-distance patterns need not be bound by arbitrary word lengths.

#### Acknowledgments

The author is grateful to: Patricia Reeder, Neil Bardhan, Carrie Miller, Kelly Johnston, Lilly Schieber, Anna States, Emily Kasman, and members of the Aslin-Newport lab. In addition, I would like to thank Jeff Heinz and two anonymous reviewers for their helpful comments and suggestions. All errors are my own. This research was supported in part by NIH grant DC000167 to E. Newport, HD37082 to R. Aslin & E. Newport, and NIH Training Grant T32DC00035.

#### References

- Albright A. Feature-based generalisation as a source of gradient acceptibility. Phonology. 2009; 26(1): 9–41.
- Archangeli, D.; Pulleyblank, D. Grounded Phonology. Cambridge: MA: The MIT Press; 1994.
- Booij G. Neutral vowels and the autosegmental analysis of Hungarian vowel harmony. Linguistics. 1984; 22:629–641.
- Cairns P, Shillcock R, Chater N, Levy J. Bootstrapping word boundaries: A bottom-up corpus-based approach to speech segmentation. Cognitive Psychology. 1997; 22:111–153. [PubMed: 9245468]
- Chelba C, Jelinek F. Structured Language Modeling. Computer Speech & Language. 2000; 14(4):283– 332.
- Chomsky, N. Lectures on government and binding. Dordecht: Foris; 1981.
- Clements GN. The autosegmental treatment of vowel harmony. Phonologica. 1976:111–119.
- Clements GN. Neutral vowels in Hungarian vowel harmony: an autosegmental interpretation. NELS. 1977; 7:49–64.
- Cohen JD, MacWhinney B, Flatt M, Provost J. PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments and Computers. 1993; 25:257–271.
- Culicover, P.; Wilkins, W. Locality in linguistic theory. New York: Academic Press; 1984.
- Finley S. The privileged status of locality in consonant harmony. Journal of Memory and Language. 2011; 65:74–83. [PubMed: 21686094]
- Finley S. Learning non-adjacent dependencies in phonology: Transparent vowels in vowel harmony. (in submission).
- Finley S, Badecker W. Analytic biases for vowel harmony languages. WCCFL 2008; 27:168–176.
- Finley S, Badecker W. Artificial grammar learning, and feature-based generalization. Journal of Memory and Language. 2009a; 61:423–437.
- Finley, S.; Badecker, W. Right-to-left biases for vowel harmony: Evidence from artificial grammar. In: Shardl, A.; Walkow, M.; Abdurrahman, M., editors. Proceedings of the 38th North East Linguistic Society Annual Meeting; 2009b. p. 269-282.
- Finley S, Badecker W. Towards a substantively biased theory of learning. BLS. 33 (in press).
- Goldsmith, J. Tone melodies and the autosegment. Paper presented at the Proceedings of the 6th Conference on African Linguistics; 1975.
- Gomez RL. Variability and detection of invariant structure. Psychological Science. 2002; 13(5):431– 436. [PubMed: 12219809]
- Hansson, GO. Ph.D. dissertation. Berkeley: University of California, Berkeley; 2001. Theoretical and typological issues in consonant harmony.
- Hayes B, Londe ZC. Stochastic phonological knowledge: The case of Hungarian vowel harmony. Phonology. 2006; 23:59–104.
- Hayes B, Wilson C. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. Linguistic Inquiry. 2008; 39:379–440.
- Heinz, J. Ph.D. dissertation. UCLA; 2007. Inductive learning of phonotactic patterns.
- Heinz J. Learning long-distance phonotactics. Linguistic Inquiry. 2010; 41(4):623-661.
- Heinz, J.; Rogers, J. Estimating strictly piecewise distributions. Proceedings of the 48th Annua Meeting of the Association for Computational Linguistics; Uppsala, Sweden: Association for Computational Linguistics; 2010.

Jurafsky, D.; Martin, JH. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. Upper Saddle River, NJ: Prentice Hall; 2008.

Keating P. Underspecification in phonetics. Phonology. 1988; 5(2):275-292.

McDonough, J. The Navajo sound system. Dordrecht: Kluwer Academic Publishers; 1990.

- Misyak, JB.; Christiansen, MH. Extending statistical learning farther and further: Long-distance dependencies, and individual differences in statistical language learning. Proceedings of the 29th Annual Conference of the Cognitive Science Society; 2007. p. 1307-1312.
- Misyak, JB.; Christiansen, MH.; Tomblin, JB. Statistical learning of nonadjacent dependencies predicts on-line processing of long-distance dependencies in natural language. Proceedings of the 31st Annual Conference of the Cognitive Science Society; 2009.
- Newport E, Aslin RN. Learning at a distance I. Statistical learning of non-adjacent dependencies. Cognitive Psychology. 2004; 48:127–162. [PubMed: 14732409]
- Onnis, L.; Christiansen, MH.; Chater, N.; Gomez, RL. Reduction of uncertainty in human sequential learning: Evidence from artificial grammar learning. Proceedings of the 25th Annual Conference of the Cognitive Science Society; 2003. p. 970-975.
- Onnis, L.; Monaghan, P.; Christiansen, MH.; Chater, N. Variability is the spice of learning, and a crucial ingredient for detection and generalizing in nonadjacent dependencies. Proceedings of the 26th Annual Conference of the Cognitive Science Society; 2004. p. 270-275.
- Pycha A, Nowak P, Shin E, Shosted R. Phonological rule-learning and its implications for a theory of vowel harmony. WCCFL 2003; 22:101–113.
- Rose S, Walker R. A typology of consonant agreement as correspondence. Language. 2004; 3:475–531.
- Sapir, E.; Hojier, H. The phonology and morphology of the Navajo language. Vol. Vol. 50. University of California Publications; 1967.
- St Clair M, Monaghan P, Ramscar M. Relationship between language structure and language learning: The suffixing preference and grammatical categorization. Cognitive Science. 2009; 33:1317–1329. [PubMed: 21585507]
- Wilson C. Experimental investigations of phonological naturalness. WCCFL. 2003; 22:101–114.
- Wilson C. Learning phonology with substantive bias: An experimental and computational study of velar palatalization. Cognitive Science. 2006; 30:945–982. [PubMed: 21702842]

# Appendix

#### Appendix 1

**Experiment 1 Stimuli** 

#### **Training Items:**

0	
∫edogi, ∫edogi∫u	sebodi, sebodisu
∫ekobi, ∫ekobi∫u	segopi, segopisu
∫epoti, ∫epoti∫u	setoko, setokosu
∫ibodu, ∫ibodu∫u	sidegu, sidegusu
∫igepu, ∫igepu∫u	sikebu, sikebusu
∫iteku, ∫iteku∫u	sipiku, sipikusu
∫obude, ∫obude∫u	soduge, sodugesu
∫ogupe, ∫ogupe∫u	sokubi, sokubisu
∫otike, ∫otike∫u	sopute, soputesu
∫udigo, ∫udigo∫u	subedo,subedosu
∫ukebo, ∫ukebo∫u	sugipo,sugiposu
∫upito, ∫huito∫u	sutiko, sutikosu

Finley

Test Items:					
Old	New			Trigram	
segopisu	*segopi∫u	seputisu	*seputi∫u	*be∫odisu	be∫odi∫u
setokosu	*setoko∫u	setokisu	*setoki∫u	busedisu	*busedi∫u
*∫edogisu	∫edogi∫u	*∫ibedusu	∫ibedu∫u	*do∫ugesu	do∫uge∫u
*∫ibodusu	∫ibodu∫u	*∫iduposu	∫idupo∫u	dosigusu	*dosigu∫u
*∫ogupesu	∫ogupe∫u	*∫okopisu	∫okopi∫u	gusepisu	*gusepi∫u
*∫otikesu	∫otike∫u	*∫otukesu	∫otuke∫u	*gu∫iposu	gu∫ipo∫u
*∫ukebosu	∫ukebo∫u	*∫ubepesu	∫ubepe∫u	*ki∫ebusu	ki∫ebu∫u
*∫upitosu	∫upito∫u	*∫ukibosu	∫ukibo∫u	kosibusu	*kosibu∫u
sidegusu	*sidegu∫u	sigopusu	*sigopu∫u	*pi∫etusu	pi∫etu∫u
sipikusu	*sipiku∫u	sipetusu	*sipetu∫u	pisotesu	*pisote∫u
sokubisu	*sokubi∫u	sobiposu	*sobipo∫u	*te∫okisu	te∫oki∫u
subedosu	*subedo∫u	sudetusu	*sudetu∫u	tesukosu	*tesuko∫u

Appendix 2

# Experiment 2 Stimuli

#### **Training Items:**

\_

besude, besudesu	bu∫ego, bu∫ego∫u
bisepu, bisepusu	di∫igu, di∫igu∫u
desoti, desotisu	di∫oge, di∫oge∫u
gosute, gosutesu	de∫ugu, de∫ugu∫u
gusiku, gusikusu	ge∫opi, ge∫opi∫u
kesobi, kesobisu	go∫ipe, go∫ipe∫u
kusigo, kusigosu	ke∫ubi, ke∫ubi∫u
pesobi, pesobisu	ko∫ide, ko∫ide∫u
piseki, pisekisu	po∫ute, po∫ute∫u
pusiko, pusikosu	pu∫eto, pu∫eto∫u
tisebu, tisebusu	ti∫oku, ti∫oku∫u
tosude, tosudesu	tu∫iko, tu∫iko∫u

Test Items:						
Old	New			New Long-Distance		ce
besudesu	*besude∫u	*be∫odisu	be∫odi∫u	seputisu	*seputi∫u	
*bu∫egosu	bu∫ego∫u	busedisu	*busedi∫u	setokisu	*setoki∫u	
desotisu	*desoti∫u	*do∫ugesu	do∫uge∫u	*∫ibedusu	∫ibedu∫u	
*de∫ugusu	de∫ugu∫u	dosigusu	*dosigu∫u	*∫iduposu	∫idupo∫u	
*ge∫opisu	ge∫opi∫u	gusepisu	*gusepi∫u	*∫okopisu	∫okopi∫u	
gusikusu	*gusiku∫u	*gu∫iposu	gu∫ipo∫u	*∫otukesu	∫otuke∫u	
kesobisu	*kesobi∫u	*ki∫ebusu	ki∫ebu∫u	*∫ubepesu	∫ubepe∫u	
*ko∫idesu	ko∫ide∫u	kosibusu	*kosibu∫u	*∫ukibosu	∫ukibo∫u	
pisekisu	*piseki∫u	*pi∫etusu	pi∫etu∫u	sigopusu	*sigopu∫u	
*po∫utesu	po∫ute∫u	pisotesu	*pisote∫u	sipetusu	*sipetu∫u	

Finley

Test Items:					
Old	New			Long-Distance	
tosudesu	*tosude∫u	*te∫okisu	te∫oki∫u	sobiposu	*sobipo∫u
*tu∫ikosu	tu∫iko∫u	tesukosu	*tesuko∫u	sudetusu	*sudetu∫u



**Figure 1.** Tiered Representations



Page 16



Finley



Page 17



Experiment 2 Results: Means (Proportion Harmonic) and Standard Errors

#### Table 1

# Examples of Exposure and Test Stimuli

	Training	Test		
Exp 1		Old	New	Trigram
Critical	∫ibodu~∫ibodu∫u	*∫ibodusu~∫ibodu∫u	*∫ukibosu~∫ukibo∫u	*be∫odisu~ be∫odi∫u
	sokubi~sokubisu	sokubisu~*sokubi∫u	sudetusu~*sudetu∫u	gusepisu~*gusepi∫u
Control	∫ibodu	See above		
	sokubi			
Exp 2		Old	New	Long-Distance
Critical	de∫ugu~de∫ugu∫u	desotisu~*desoti∫u	*be∫odisu~ be∫odi∫u	*∫ukibosu~∫ukibo∫u
	po∫ute~po∫ute∫u	*po∫utesu~poshute∫u	gusepisu~*gusep i∫u	sudetusu~*sudetu∫u
Control	de∫ugu		See above	
	po∫ute			