LEARNING CONSONANT HARMONY IN ARTIFICAL LANGUAGES*

Kevin McMullin
University of British Columbia

1. Introduction

1.1 Learning bias in phonology

A great deal of research in linguistics concentrates on identifying, describing, and analyzing phonological patterns. However, we still do not know how humans learn them. This paper focuses on the learning of one such pattern, consonant harmony, in which two non-adjacent consonants in are required to agree in some way. In the Bantu language Yaka, for example, there is a perfective suffix *-ili*, which attaches to a verb root. However, if the root contains a nasal consonant, then the [l] in the suffix becomes a nasal [n]. Thus, the verb stem *ja**n**-a* 'cry out in pain' is *ja**n**-i**n**i* in the perfective rather than **ja**n**-ili* (Hyman 1995). More than 130 languages have some form of consonant harmony system (Hansson 2010a), each with its own set of properties. While some of these properties are quite common, others are rare of unattested, but we do not know why. One possible explanation is that humans have certain cognitive learning biases that affect language learnability; the rare patterns are simply harder to learn, so they are less likely to ever arise in a language, let alone persist over time (see Rafferty *et al.* 2013 for limitations). Current research in linguistics and cognitive psychology has lent support to this idea by showing that some patterns involving the interaction of non-adjacent sounds are indeed more difficult to learn than others (Creel *et al*. 2004, Newport & Aslin 2004), particularly with respect to the relative similarity between the sounds (Gebhart *et al*. 2009, Moreton 2012). This generalization is mirrored in the typology of consonant harmony, as two similar consonants like [s] and [ʃ], or [l] and [r] are much more likely to interact than two dissimilar consonants such as [m] and [k], or [s] and [b]. The present study extends beyond the issue of similarity and investigates the learnability of consonant harmony with respect to locality, hypothesizing that humans will learn and generalize consonant harmony patterns in a way that matches their typological distribution.

1.2 Typology of consonant harmony: locality

This section provides a brief description of the typology of consonant harmony patterns, focusing especially on generalizations that can be made about locality

* The research reported here was supported by SSHRC Standard Research Grant 410-2009-2831, awarded to Gunnar Ólafur Hansson. Many thanks to Gunnar Ólafur Hansson, Carla Hudson Kam, and Kathleen Currie Hall for discussion and feedback, as well as to James Crippen for his help in creating the experimental stimuli.

Actes du congrès annuel de l'Association canadienne de linguistique 2013.
Proceedings of the 2013 annual conference of the Canadian Linguistic Association.

in terms of the relative distance between two interacting consonants (see Hansson 2001, 2010a, Rose & Walker 2004 for comprehensive typological surveys of consonant harmony). Languages that have a consonant harmony system require two interacting consonants to agree with respect to some feature. The relevant feature varies by language and can be voicing, nasality, palatalization, and so on. The most common type of consonant harmony is sibilant harmony, which involves an interaction between sibilant consonants such as {s,ʃ,z,ʒ,ts,dz,tʃ,dʒ}. For example, the Samala (Inseño Chumash) language of Southern California prohibits the co-occurrence of sibilants that do not match in anteriority, disallowing sequences of s...ʃ and ʃ...s within a word, no matter the distance between them (Applegate 1972; Hansson 2010a), as illustrated in (1).

(1)	a.	/ha-**s**-xintila/	→	[ha**s**xintila]	'his gentile name'
	b.	/ha-**s**-xintila-wa**ʃ**/	→	[ha**ʃ**xintilawa**ʃ**]	'his former gentile name'

The typology of consonant harmony reveals an interesting split with respect to locality, resulting in just two types of languages. First, there are languages that apply harmony only when the interacting consonants are in a transvocalic relationship, separated by at most one vowel. Second, there are languages that apply harmony anytime the two consonants co-occur in the relevant domain (e.g. within a word), no matter the distance between them. The attested split is illustrated in (2) and (3) with sibilant harmony in two Omotic languages, Koyra and Aari.

(2) Koyra: transvocalic harmony (Hayward 1982)

	a.	/ʔor**dʒ**-u**s**-/	→	[ʔor**dʒ**-u**ʃ**-]	'make big, increate (tr.)'
	b.	/ʃod-us-/	→	[ʃod-us-] (*[ʃoduʃ])	'cause to uproot'

(3) Aari: unbounded harmony (Hayward 1990)

	a.	/ʔu**ʃ**-**s**i**s**-/	→	[ʔu**ʃ**-**ʃ**i**ʃ**-]	'cause to cook'
	b.	/**ʃ**ed-er-**s**-it/	→	[**ʃ**ed-er-**ʃ**-it]	'I was seen'

It is not clear why languages allow consonant harmony to apply across either transvocalic or unbounded distances, but nothing else. For example, there is no language that applies harmony when the two consonants are separated by one or two vowels but not more, and there is no language that applies harmony everywhere except in transvocalic contexts. This paper presents experimental evidence in support of the hypothesis that this dichotomy is the result of a learning bias that allows humans to discover only these two types of patterns.

1.3 Artificial language learning

The traditional way to study phonological learning is to observe infants as they acquire the patterns in their native language. However, as many phonological patterns are relatively rare, access to children learning the language is not always feasible. As a result, it is becoming increasingly common for researchers to construct artificial languages for experimental studies (e.g. Finley & Badecker 2009, Moreton 2012, Nevins 2010). In a typical experiment, subjects complete a training phase in which they are exposed words from an artificial language that exhibits some pattern, followed by a testing phase to determine what they have learned. Both of these phases can take many different forms. In the testing phase, for example, some experiments use a 2 alternative forced choice task in which the subject must decide which of two words is "correct" or "sounds better". Alternatively, subjects may be asked to produce certain words, and their responses as well as the errors they make can be taken as evidence for what they have or have not learned (e.g. Rose and King 2007).

This methodology creates an accessible way to study any type of phonological pattern, whether it is relatively frequent, rare, or completely unattested across the world's languages. An additional benefit of this method is that it allows the researcher to have control over the input that the learner gets, both in terms of the amount of input and its content. These are the obvious advantages for those interested in questions about how humans learn, but the methodology is not without criticisms (see Moreton & Pater 2012a,b for an overview of the findings and criticisms of these experiments). For example, we may not know what biases the learners are bringing in from their own native language, or their language experience as a whole. Furthermore, we still do not know about the relationship between how children and adults learn the patterns in a new language. These are ongoing, unresolved debates, but having a well-constructed control condition can help in understanding what biases a learner might come in with, so that we can build them into the statistical models used to analyze experimental results.

1.4 Previous studies

Of particular interest here are studies that have investigated the learning of consonant harmony at different levels of locality. Finley (2011) examined how adults learn transvocalic dependencies in an artificial language with sibilant harmony, as compared with nonlocal dependencies (across two vowels). In the training phase of experiment 1, subjects heard pairs of words. The first word was a two syllable stem of the form cvSv, where S is a sibilant [s] or [ʃ]. The second was the same stem with a suffix -Su, where the sibilant in the suffix was identical to the sibilant in the stem. Thus, all suffixed training items showed evidence of "first-order" harmony. In the testing phase, subjects chose which of two suffixed forms was correct, one of which had matching sibilants (harmony) while the other had mismatched sibilants. Experiment 2 was similar, but learners were trained only on a "second-order" pattern, in which the stems were of the form Svcv. The results indicated an asymmetry that reflects the typology described above. That is, subjects in experiment 1 learned the first-order pattern for both familiar and unfamiliar first-order test items, and did not generalize this to the second-order contexts, in which the sibilants were further apart. Subjects in experiment 2 were successful in learning the second-order pattern and they generalized the pattern inward to the first-order context. The author used these results as evidence for first-order patterns having a 'privileged status'.

Finley (2012) expands the study by presenting subjects with longer stems to investigate the role of distance between two interacting consonants. Subjects were trained on pairs of words consisting of a three-syllable stem with one sibilant followed by the suffixed form, in which the sibilant in the suffix -Su displayed harmony with the stem. This study makes two important conclusions. First, that learners are able to learn this type of sibilant harmony pattern even when the sibilants are up to five segments away (i.e. Svcvcv-Su); the further away the dependency, the harder it is to learn. Second, when learners are trained on a "second-order" pattern, with cvSvcv-Su words, they generalize to even longer distances: Svcvcv-Su. Based on these results, Finley argues that when learners are tasked with discovering a long-distance interaction, they do so without reference to intervening distance, and so they apply the pattern to all contexts. Crucially, this generalization excludes the local-only patterns that apply harmony when the distance is at most transvocalic.

1.5 The present study

The remainder of this paper presents a replication of Finley's findings using a different, and arguably improved methodology, described in Section 2, with a more appropriate statistical analysis of the results, presented in Section 3. In Section 4, I discuss the findings as evidence for a cognitive learning bias and argue that such biases can influence the typological shape of the world's phonological patterns, in this case consonant harmony. Furthermore, I outline a computational learning algorithm, a precedence learner (Heinz 2010), that can potentially account for this asymmetry. Section 5 provides a summary of the study and concluding remarks.

2. Methodology

2.1 Subjects

Thirty-six adults aged 18-40 participated in the study (25 female, 8 male, 3 unspecified; mean age of 24). All were native speakers of North American English, reported no speech or hearing disorders and had no experience with a language containing a harmony system. Subjects were compensated $10 for participating in the experiment, which took about 45 minutes to complete.

2.2 Stimuli

All stimuli were recorded by a phonetically trained, male native English speaker. While he knew that the stimuli would be used for an artificial language experiment, he was unaware of the exact topic of study and the hypothesis. Most of the stimuli consisted of three-syllable "verbs" which took the form CVCVCV. There were also six two-syllable verbs of the form CVCV for the practice phase, as described below. Additionally, the same speaker recorded two corresponding suffixed verbs for each root, one of each -su and -ʃi. Consonants in the stems were chosen from sibilants [s, ʃ] as well as stop consonants [p,t,k,b,d,g] and vowels were chosen from [i,e,a,o,u]. The stimuli set was carefully constructed such that there was no other predictable pattern in the data. Crucially, the same number of each vowel and consonant appeared in each position for the verb roots containing [s] or [ʃ].

2.3 Design and procedure

Subjects were told that their task was to practice pronouncing words from a new language and to learn how to say verbs from the language in the past and future tense. Each subject worked through three phases of the experiment.

In the **practice phase**, subjects were told how to conjugate verbs in the language. They then heard, over a set of headphones, six pairs of words for the past tense (a verb root followed by a suffixed form with -su), and six word pairs for the future tense (this time with the suffix -ʃi). The six verb roots in this portion were just two syllables and contained no sibilants. As a result, there was no input with any evidence of an alternation in the practice phase.

In the **training phase**, subjects heard triplets of words, the first of which was always a three-syllable verb root. Half of these contained one sibilant (one quarter with [s], one quarter with [ʃ]) and the other half contained only stop consonants. This was done in part to test how learners perform without 100% of the input containing meaningful information, and in part to allow the other half of the input to be manipulated in future studies. The two subsequent words in each triplet were the four-syllable suffixed forms. Since both suffixes begin with a sibilant, if the verb root also contained a sibilant, one of the suffixed forms would exhibit suffix-triggered harmony. For example, given the verb *bugaso*, the suffixed forms would be *bugaso-su* and *bugaʃo-ʃi*. Each triplet was presented twice, with the suffixed forms counterbalanced for order and the order of triplets was randomized for each subject. The subjects were asked to repeat each word into a head-mounted microphone that recorded their productions. This was done to allow for possible further analysis, as well as to reinforce the training process. In total, each subject heard and repeated 120 triplets twice each for a total of 720 productions. The words in both the practice phase and the training phase were presented over the headphones only, and the subjects got no information about any semantic content of the words, except that they were verbs that can be conjugated in these two tenses. Subjects were assigned to one of three training conditions, described below in section 2.4, that differed in the types of words presented in their training phase.

In the **testing phase**, subjects heard 30 new verbs, and then completed a forced choice task, in which they were asked to choose the correct conjugated form of the verb. Testing items included 30 suffixed harmonic/non-harmonic verb pairs of the form cvcvcv-Sv. These pairs consisted of ten each with the sibilant in the three different consonantal positions in the root. There were an equal number of -su and -ʃi forms, and the order of presentation was randomized for each subject. It is important to note that all testing items required the subject to choose an alternation in the root in order to identify the suffixed form with harmony (i.e. if the testing root contained a [s], then the corresponding suffixed forms would both have the suffix -ʃi). It would be ideal to have many filler items in the testing phase, as well as harmonic responses that do not require an alternation. However, a confound of that approach is that any items presented in the testing phase effectively become a part of the training, so it is necessary to both limit the testing items and maximize the amount of relevant data from each subject. As a result, subjects were tested only on cases that required the choice of an unfaithful root alternation in favour of harmony with the sibilant in the suffix. Subjects were given a maximum of 3 seconds after the onset of the final word in the triplet to register a response, and were presented with their response time after each successful trial.

2.4 Training conditions

The subjects were divided into three conditions: *Local*, *Nonlocal*, and *Control*. Stimuli for these three groups differed only in the training phase, with the testing items being the same for all groups. For the portion of the training items that contained sibilants, the *Local* group was trained on verb roots that contained sibilants only in the last consonant position (i.e. cvcvSv), while the *Nonlocal* group was trained on cvSvcv roots. The *Control* group was trained only on verbs that contained no sibilants. This was done to reveal any potential pre-existing biases that English speakers may have toward choosing harmony, or any biases introduced by the experimental procedure.

3. Results and analysis

As reported in section 2.1, thirty-six subjects participated in this study. However, the data from six subjects was dropped as the result of failure to follow instructions (2 subjects), choosing either the first or second test item on all recorded trials (2 subjects), or equipment failure (2 subjects). Of the remaining 30 subjects, 10 were in each condition and 827 of a possible 900 responses were registered within the allotted three-second timeframe.

3.1 An overview of responses

The first step in analyzing the results is to get a clear picture of what subjects' choices were in each of the training conditions and for each type of testing item. To do this, I will present the results in this section as the proportion of testing items in which subjects chose to apply consonant harmony rather than stay faithful to the root. Keeping in mind that subjects in the control condition saw no evidence of an alternation in their input, their scores should be close to 0% if they are in no way biased towards harmony. Figure 1 on the following page presents the proportion of responses where the subject chose a non-faithful alternation in the root in order to have harmony with the sibilant in the suffix. The results here are intended to give an overall picture of what the subjects in each condition chose, but will not be used to test for statistical significance.

The results shown in Figure 1 can be summarized as follows. Subjects in the control group are most likely to choose an alternation that results in harmony for sibilants in the second syllable of the root, and are least likely to choose an alternation for word-initial sibilants. Compared to the control group, subjects who were trained locally are more likely to choose an alternation in all positions, though this is most evident in the local cases, followed by the cases with a sibilant in the second syllable, with minimal distinction in the word initial position (at distance 3). Subjects in the Nonlocal training condition are also more likely to choose an alternation at all distances compared to the control group. Compared to the local group, they are less likely to do so at distance 1, but more likely at distances 2 and 3.

3.2 A mixed-effects logistic regression model

Rather than using the mean proportions for each subject, a more appropriate way to analyze categorical data is with a logistic regression model (see Jaeger 2008). In a logistic regression, the model finds the best fit for the log odds of choosing one response or another, in this case harmony vs. no harmony, based on a set of predictor variables. The predictor variables here, which are also categorical values, are the training conditions (including Control, Local, Nonlocal), as well as the distance between sibilants for each testing item (1, 2, or 3). Note that the

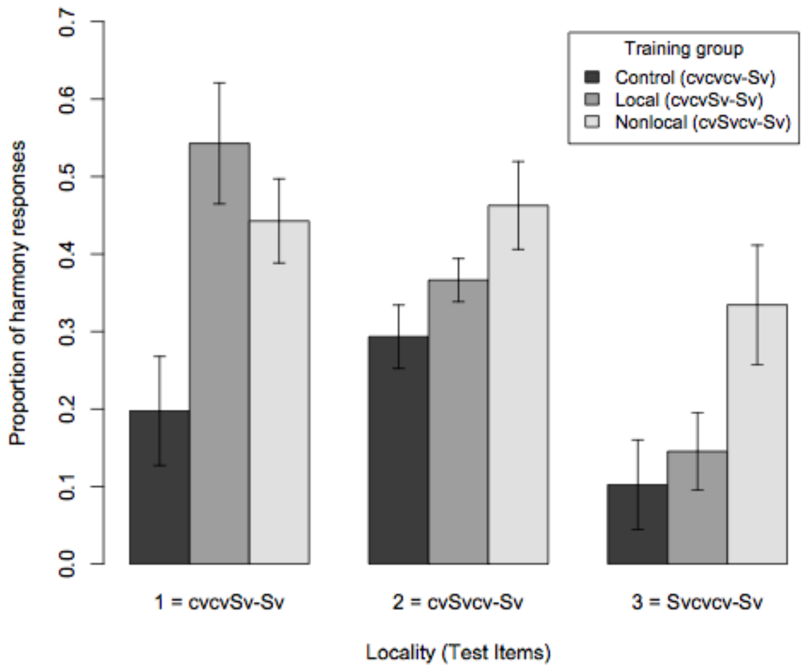

Figure 1 – Proportion of harmony responses at each testing distance for each training condition. Error bars represent standard error, based on the means of subjects in that group.

hypothesis predicts different results for each group at each distance, so it is important to include interactions between group and distance in the model as well. Other variables (including trial number, whether the test item with harmony was presented first or second, and whether the triggering suffix was -ʃi or -su) were included as fixed effects in other models, but showed no significant effects and the models including these variables did not give a significantly better fit to the data, and so are omitted from the model. Additionally, with the use of a mixed-effects logistic regression, the model can account for tendencies for individual subjects. Barr *et al.* (2013) argues for a maximally rich random-effects structure in regression, so long as it is justified by the design. In the model presented below, a random intercept as well as a random slope for whether the harmony item was presented first or second is included for each subject. The model given in Table 1 was created with the *glmer* function from the *lme4* package (Bates *et al.* 2012) implemented in R (R Core Team 2012).

In Table 1, the estimate for the Intercept indicates the model's guess for the likelihood (in terms of log odds) that an average subject in the control group will choose harmony when presented with a distance 1 testing item. The negative estimate of -1.504 indicates that the model predicts the odds of choosing harmony to be exp(-1.504)=0.222, which, in terms of probability, translates to a 18.2% chance. The estimates for the subsequent main effects of predictor variables indicate whether there is an increase or decrease in the likelihood of choosing harmony, as compared to this baseline.

Table 1 – A summary of the fixed effects portion of the mixed-effects logistic regression model
(*N*=827, 30 Subjects, log-likelihood -460.6)

Coefficient	Estimate	SE	Pr(>|z|)
Intercept	-1.504	0.336	<0.001*
Distance 2	0.613	0.369	0.097
Distance 3	-0.794	0.446	0.075
Local Training	1.686	0.443	<0.001*
Nonlocal Training	1.160	0.445	0.009*
Distance 2 x Local	-1.383	0.480	0.003*
Distance 3 x Local	-1.308	0.582	0.025*
Distance 2 x Nonlocal	-0.549	0.484	0.257
Distance 3 x Nonlocal	0.257	0.547	0.638

With respect to the main effects of testing distance, the estimate for Distance 2 is positive, indicating an increase in the likelihood of Control subjects choosing harmony, while the estimate for Distance 3 is negative, indicating a decrease, though neither of these effects reaches a significance level of <0.05. For the main effects of training group, both the Local and Nonlocal groups have positive estimates that are significant, and so they are much more likely to choose harmony at Distance 1. This indicates that they have learned a consonant harmony pattern, or at least have begun to learn it. For the local group, this is the pattern they were trained on, but for the Nonlocal group, this effect demonstrates that they are choosing harmony at Distance 1, even though they were trained at Distance 2 and never saw evidence for what to do for sibilants in a transvocalic relationship.

For the interactions of distance and group, the estimates indicate whether the likelihood of choosing harmony increases or decreases after already taking into account the main effects of group and distance. For the Local group, there is a significant decrease in the likelihood of choosing harmony at both Distance 2 and Distance 3. This indicates that they have not generalized the local pattern that they learned to either of the nonlocal distances. In contrast, the interactions of the Nonlocal group with both Distances 2 and 3 did not result in a significant increase or decrease in the likelihood of choosing harmony indicating that the pattern they have learned applies to all distances.

4. Discussion

4.1 Comparing results to typology

The results presented above provide evidence in support of the hypothesis that humans will learn new phonological patterns and generalize them in a way that matches up to the typology of the world's languages. As described in Section 1, the typology of consonant harmony reveals just two types of languages with respect to locality. The first type consists of the languages that apply harmony only when the two dependent consonants are separated by at most one vowel. The second type is the set of languages that apply harmony whenever the relevant consonants co-occur within some domain, as defined without reference to any phonological representation, including transvocalic contexts as well as longer distances. In the results of this experiment, we see that humans learning a new artificial language can learn both types of languages. However, it is how the learners tend to generalize the pattern that makes this result most interesting.

Consider the learners in the Local group. Their relevant training items were restricted to instances of harmony applying in transvocalic contexts. Not only were they more likely than the Control group to choose a harmonic response in the same types of items that they were trained on, but they did not generalize this to nonlocal distances. Once the learners in the Local group discover the pattern, they do not extend it to nonlocal distances, even though such a pattern is attested in natural languages, and they saw no evidence to the contrary. That is, the subjects seem to be making the most direct application of their training as is possible – harmony occurs in local contexts, and nowhere else.

We now turn to the Nonlocal group, whose relevant training items were restricted to instances of nonlocal harmony applying over two syllables (cvSvcv-Sv). Subjects in this group were more likely than the Control group to apply harmony in the same types of words that they were trained on. Furthermore, they generalized this pattern to types of words that were not encountered in their training data – to roots that had a sibilant in the final syllable (a shorter distance 1), and roots that had a sibilant in the first syllable (a longer distance 3). So in this case, subjects discovered a pattern that applied harmony at a nonlocal distance. Additionally, they correctly rule out the possibility of it being the first type of language that applies harmony only at local distances, and so they settle on generalizing it to all contexts within the word, in line with the pattern we see in the second, and only other type of attested consonant harmony language.

4.2 A learning bias

An increasingly common explanation for why some phonological patterns are more common than others is that the more frequent patterns are more likely to survive over time because of learning biases. The results discussed above suggest that humans have an analytic learning bias that facilitates the learning of some consonant harmony patterns, prevents the learning of others, and helps determine how a pattern is generalized. Over time, a propensity for humans to learn a certain pattern would result in the pattern having a high survival rate (Moreton 2008). The subjects, who had no experience with a language containing harmony, were able to learn the pattern and they generalized in a way that matched the typology. This is taken as evidence that it is a learning bias that has resulted in only two types of locality in the consonant harmony patterns found in the world's languages.

4.3 Formal models and other possible outcomes

The results of this experiment support the hypothesis that humans will generalize a phonological pattern in a way that mirrors the typology of that pattern. However, note that the subjects in this experiment spoke no languages other than English, and presumably know nothing about the typology of consonant harmony systems. This section then, will ignore the typological facts and consider the possible outcomes from a learner that has no preconceived notion of how consonant harmony patterns should apply.

First, consider a learner that is trained only on items that exhibit local harmony, and some viable strategies that the learner might use to discover a consonant harmony pattern. One possible type of learner is one that relies on brute force, looking for a dependency between any two consonants in a word, no matter how far apart they are and without keeping track of distances. This learner, though trained only on harmony in local contexts, will learn only that there is a dependency between [s] and [ʃ]. When presented with testing items, it will incorrectly apply this pattern to both local and nonlocal sibilant pairs, since it does not record the distance. In order to prohibit it from overextending the pattern, there are at least two possible modifications that can be made. First is to allow the learner to also keep track of the distance between the two consonants. In the training phase, the learner will discover that a dependency exists between sibilants that are separated by a vowel. In the testing phase, it will look for any such instances and apply harmony. A second, similar option is to have the learner restrict its maximum search space to the largest distance between sibilants seen in the training data. It will have completed the training phase having never seen any instances of two sibilants that are separated by more than one vowel. During the testing phase, it will not even look for sibilants in a nonlocal relationship, and so it will apply harmony in the local cases only. Either of the latter two learners would give a result analogous to what we observed in the experiment for the local group.

With this in mind, however, a problem arises when considering the task of learning the pattern with input restricted to harmony at distance 2. Only the first of the three learners presented above, which incorrectly generalized a local pattern to nonlocal contexts, is capable of learning and generalizing in the same way as the nonlocal group. In this case, the modifications that solved this problem for the case of local harmony would result in it learning to apply harmony either at distance 2 only, or to apply it at any distance up to 2. This paradox emphasizes the fact that even though human learners tend to generalize consonant harmony patterns in a way that directly relates to the typology, finding one learning algorithm that can simultaneously learn both types of languages is not trivial.

4.4 The precedence model of learning long-distance dependencies

In an attempt to explain how humans might learn long-distance dependencies, including consonant harmony, Heinz (2010) proposes a precedence model of learning long-distance phonotactics. In principle, it is a very simple model, and in practice it makes strong predictions about what the typology of consonant harmony should look like if language learners use the proposed strategy when learning the pattern.

In the precedence model of learning, a learner takes into account *precedence relationships* in addition to keeping track of bigrams. For example, in a word like “pants” /pænts/, the learner keeps track of the bigrams {p,æ},{æ,n},{n,t}, and {t,s}, which are all adjacent phonemes. Additionally, the learner tracks the precedence relationships {p,æ},{p,n},{p,t},{p,s},{æ,n},{æ,t}, {æ,s},{n,t},{n,s} and {t,s} without reference to distance, just that the first sound precedes the second somewhere in the word. Therefore, such a learner makes a distinction between co-occurrence restrictions between adjacent segments, and dependencies that hold between any two segments. For a case of nonlocal sibilant harmony that is purely phonotactic, the precedence relations {s,ʃ} and {ʃ,s} would never be encountered in the language, so the learner would recognize that two different sibilants should no co-occur within a word. Note that this model does not keep track of any segments that occur in between the segments of each precedence relationship, and thus no intervening segment can have any influence on the nature of the precedence relationship. Such a property makes the strong prediction that there is no possibility of an intervening segment blocking the interaction between non-adjacent phonemes, or at least that such a property would never be learned. Heinz (2010) argues that this property is desirable since there are no discovered instances of blocking in the typology of consonant harmony. However, several languages, including some Berber dialects (Elmedlaoui 1995, Hansson 2010b) and Kinyarwanda (Walker & Mpiranya 2005), may exhibit the blocking of consonant harmony, so a reevaluation of this aspect of the precedence learning model is in order.

There is another prediction that the precedence model of learning makes by adding the notion of a consonant tier. By doing so, the learner can keep track of bigram and precedence relationships between consonants alone, therefore making a distinction between consonants that are adjacent on the consonant tier, and consonants that are in precedence relations, but ignoring any further distinction regarding distance. This would give rise to the typological split mentioned above between languages that have only local dependencies, and languages in which the dependency holds at all distances. One downside to this model is that it does not account for the effects of similarity. The learner is equally capable of picking out dependencies among sounds that are relatively dissimilar, which is reflected neither in the typology of consonant harmony, nor in the experimental research reported in the literature. However, this algorithm is not incompatible with other models of learning that are biased towards picking out similar segments (e.g. Hayes & Wilson 2008), and such a combination could provide a more comprehensive model of how humans are learning phonological patterns.

5. Summary and conclusions

The goal of this study was to give an explanation for why the typology of consonant harmony includes only two types of languages with respect to locality – those with transvocalic harmony, and those with unbounded harmony that applies across all distances. A mixed-effects logistic regression model showed that learners trained on sibilant harmony only at a local (transvocalic) distance learned the pattern, but did not generalize the pattern to sibilants at nonlocal distances (two or three syllables away from the triggering suffix). Learners trained on sibilant harmony only at the nonlocal distance (two syllables away) learned this pattern, and generalized not only to the shorter, local distance, but also to the word initial sibilants that were three syllables away from the triggering suffix. This experiment, with several differences in methodology and data analysis, has replicated findings from Finley (2011, 2012) and increases the evidence that there are learning biases affecting how humans learn and generalize, and that these biases can influence the typological shape of linguistic patterns.

References

Applegate, Richard. 1972. *Inseño Chumash grammar*. Doctoral dissertation, University of California, Berkeley.

Barr, Douglas, Roger Levy, Cristoph Scheepers, and Harry Tily. 2013. Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68:255-278.

Bates, Douglas, Martin Maechler, and Ben Bolker. 2012. lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4.

Creel, Sarah, Elissa Newport, and Richard Aslin. 2004. Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning Memory, and Cognition* 30:1119-1130.

Elmedlaoui, Mohamed. 1995. *Aspects des respesentations phonologiques dans certains languages chamito-semitiques*. Rabat: Faculté des Lettres et des Sciences Humaines. [Doctoral dissertation, Université Mohammed V, 1992.]

Finley, Sara. 2011. The privileged status of locality in consonant harmony. *Journal of Memory and Language* 65:74-83.

Finley, Sara. 2012. Testing the limits of long-distance learning: Learning beyond the three-segment window. *Cognitive Science* 36:740-756.

Finley, Sara, and William Badecker. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language* 61:423-437.

Hansson, Gunnar Ólafur. 2001. *Theoretical and typological issues in consonant harmony*. Doctoral dissertation, University of California, Berkeley.

Hansson, Gunnar Ólafur. 2010a. *Consonant harmony: long-distance interaction in phonology*. Berkeley, CA: University of California Press.

Hansson, Gunnar Ólafur. 2010b. Long-distance voicing assimilation in Berber: spreading and/or agreement? *Actes du Congrès de l'ACL 2010 / 2010 CLA Conference Proceedings*, ed. by M. Heijl. Association canadienne de linguistique / Canadian Linguistic Association.

Hayes, Bruce, and Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39:379-440.

Hayward, Richard. 1982. Notes on the Koyra language. *Afrika and Überese* 65:211-268.

Hayward, Richard. 1990. Notes on the Aari language. In *Omotic Language Studies* ed. Richard Hayward. School of Oriental and African Studies, University of London.

Heinz, Jeffrey. 2010. Learning long-distance phonotactics. *Linguistic Inquiry* 41:623-661.

Hyman, Larry. 1995. Nasal consonant harmony at a distance: the case of Yaka. *Studies in African Linguistics* 24:5-30.

Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language,* 59: 434-446.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25:83-127.

Moreton, Elliott. 2012. Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language* 67:165-183.

Moreton, Elliott, and Joe Pater. 2012a. Structure and substance in artificial-phonology learning, Part I: Structure. *Language & Linguistics Compass* 6:686-701.

Moreton, Elliott, and Joe Pater. 2012b. Structure and substance in artificial-phonology learning, Part II: Substance. *Language & Linguistics Compass* 6:702-718.

Nevins, Andrew. 2010. Two case studies in phonological universals: a view from artificial grammars. *Biolinguistics* 4: 218-233.

Newport, Elissa, and Richard Aslin. 2004. Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology* 48:127-162.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rafferty, Anna, Thomas Griffiths, and Marc Ettlinger. 2013. Greater learnability is not sufficient to produce cultural universals. *Cognition* 129:70-87.

Rose, Sharon, and Lisa King. 2007. Speech error elicitation and co-occurrence restrictions in two Ethiopian Semitic languages. *Language and Speech* 50:451-504.

Rose, Sharon, and Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80:475-531.

Walker, Rachel, and Fidèle Mpiranya. 2005. On triggers and opacity in coronal harmony. *Proceedings of BLS* 31: 383–394.