Typological Asymmetries in Round Vowel Harmony: Support from Artificial Grammar Learning

Sara Finley
Department of Psychology Elmhurst College 190 Prospect Ave Elmhurst, IL 60126
finleys@elmhurst.edu

Abstract

Providing evidence for the universal tendencies of patterns in the world's languages can be difficult, as it is impossible to sample all possible languages, and linguistic samples are subject to interpretation. However, experimental techniques such as artificial grammar learning paradigms make it possible to uncover the psychological reality of claimed universal tendencies. This paper addresses learning of phonological patterns (systematic tendencies in the sounds in language). Specifically, I explore the role of phonetic grounding in learning round harmony, a phonological process in which words must contain either all round vowels ([o, u]) or all unround vowels ([i, e]). The phonetic precursors to round harmony are such that mid vowels ([o, e]), which receive the greatest perceptual benefit from harmony, are most likely to trigger harmony. High vowels ([i, u]), however, are cross-linguistically less likely to trigger round harmony. Adult participants were exposed to a miniature language that contained a round harmony pattern in which the harmony source triggers were either high vowels ([i, u]) (poor harmony source triggers) or mid vowels ([o, e]) (ideal harmony source triggers). Only participants who were exposed to the ideal mid vowel harmony source triggers were successfully able to generalize the harmony pattern to novel instances, suggesting that perception and phonetic naturalness play a role in learning.

Linguistic universals have been proposed as a way of understanding the mechanisms that underlie the structure of language, and how language might be learned. However, achieving a full understanding of linguistic universals is challenging because one can never be absolutely certain that a 'universal' tendency is truly universal (Evans & Levinson, 2009). It is impossible to study every language that exists, has existed, or will exist in the future. Thus, universals are based on an incomplete sample of languages. Because one must take into account historical factors in making conclusions about universal tendencies, even an a-theoretical descriptive analysis can be subject to debate. Additionally, there is still of a question of the psychological validity of proposed universals. A linguistic universal may not necessarily be part of the cognitive abilities of the language learner, but may be the result of a unrelated variables, or may be a result of a series of historical changes (which may or may not be grounded in cognitive or phonetic principles) (Blevins, 2004).

Artificial grammar learning paradigms may offer a potential solution to the problem of uncovering linguistic universals (Nevins, 2009). In an artificial grammar learning paradigm, researchers can control for factors that are otherwise uncontrollable (Finley & Badecker, 2008; Guest, Dell, & Cole, 2000). It is possible to compare minimally different languages that are believed to follow universal tendencies with languages that are not believed to follow such tendencies. The present paper presents evidence that typological tendencies in phonology are related to learning biases. Specifically, I address the role of phonetic

naturalness in phonological patterns. Phonological patterns are derived from systematic rules (or constraints) that govern the distribution of sounds and sound sequences.

While both phonetically grounded and phonetically ungrounded patterns exist in natural language (Anderson, 1981), there are proposed universal tendencies for phonological patterns. Because languages (specifically English) do not make use of all phonological patterns, it is possible to expose naïve (English speaking) participants to novel phonological patterns and test whether such participants are more likely to learn the pattern that follows a proposed universal tendency. Even in an artificial setting, assessing the potential advantage of learning a phonetically natural pattern is difficult because phonetic naturalness and complexity are often conflated. Further, different phonological patterns may be learned at different rates independently of phonetic naturalness (e.g., based on regularity or frequency of occurrence in a given language). The present study works around these challenges by comparing the learnability of natural and unnatural variants of a single phonological process (round vowel harmony). Vowel harmony is a phonological pattern in which vowels must share the same value of a phonological feature (e.g., backness, round). In Turkish[1], which has back and round harmony, the back and round features of the stem vowel determine the back and round features of suffix vowels (e.g., [pul-**u**n] 'stamp', [ip-**i**n] 'rope') (Clements & Sezer, 1982).

For the purpose of this paper, the term 'source triggers' refers to vowels that initiate a harmonic domain, while 'target undergoers' refers to vowels that undergo harmony. The term 'participating vowel' refers to any vowel that undergoes or initiates harmony, while 'non-participating vowel' refers to any vowel that is neither a source trigger or a target undergoer for harmony.

There are several possible universal factors that influence whether vowel harmony applies, including inventory constraints, coarticulation and perceptual prominence (Kaun, 2004; Korn, 1969; Suomi, 1983; Walker, 2005). For example, cross-linguistic studies of round harmony have shown that the vowels that receive the greatest perceptual benefit from harmony are most likely to be source triggers for round harmony (Suomi, 1983; Walker, 2005)[2]. Round harmony increases the perceptibility of the round feature (Kaun, 1995, 2004), which subsequently increases the likelihood that a round vowel will be accurately identified as round. The impact of this increase in perceptibility of rounding varies depending on the height of the harmony source trigger. Mid vowels tend to show a greater perceptual benefit for vowel harmony compared to high vowels (Terbeek, 1977).

The preference for mid over high vowels as source triggers for round harmony may also reflect the fact that it is easier to identify the round feature for high vowels compared to mid vowels (Kaun, 1995). Thus, mid vowels receive more perceptual benefit from harmony than high vowels. Using multidimensional scaling, Terbeek (1977) created a continuum of the perceptibility of the round feature. In this continuum, high and back vowels are most likely to be perceived as round, while mid and front vowels are less likely to be perceived as round. In addition, Linker (1982) created a rounding continuum that was based on articulatory measurements (horizontal opening and lip opening) in several languages (including English). Linker demonstrated that high vowels are produced as more 'canonically round' than mid vowels. These facts help explain the typological asymmetry explored by Kaun (1995) and Korn (1969); mid vowels are less likely to be restricted as harmony source triggers than high vowels.

---

[1]This example is simplified. The reader is directed to Clements and Sezer (1982) for a complete description of Turkish harmony.
[2]Such asymmetries have also been shown for height harmony (Walker, 2005), but are not necessarily present in all types of harmony. The present study therefore addresses the notion of typologically salient triggers in learning for round harmony only.

Back vowel harmony requires agreement in back features (e.g., front vowels /i, e/ cannot co-occur with back vowels /o, u/), and typically co-occurs with round harmony (e.g., Turkish). Asymmetries between mid and high vowels as source triggers for harmony can also be found in back harmony. For example, if the mid front vowel [e] does not trigger harmony, the high vowel [i] will also fail to trigger harmony (van der Hulst & van de Weijer, 1995). In Finnish, both [i] and [e] are neutral to back harmony, but diachronically, [i] was the first to become neutral (Suomi 1983). While Suomi (1983) argues that neutrality in back harmony is typically based on the perceptibility of target undergoers, the perceptually grounded constraint against high vowels as harmony triggers may indirectly contribute to an overall preference for mid vowels over high vowels as triggers of round harmony. Thus, the constraint that favors /e/ as a trigger of back harmony may contribute to a bias for /e/ over /i/ in languages with both round and back harmony.

The cross-linguistic evidence supports the notion of a constraint that prefers mid vowels as source triggers for round harmony. While there is phonetic evidence to support this claim, there is also reason to be skeptical. Vowel harmony does not occur in all languages, and round harmony is found in only a subset of languages that show some form of harmony. In addition, many languages with round harmony are related in some way (e.g., Turkic languages (Korn, 1969)), raising the possibility that restrictions on high vowel source triggers in round harmony are due to historical changes unrelated to phonetic biases. Further, because round harmony languages contain mid vowel harmony source triggers, it is unclear whether mid vowel source triggers are necessary to learn a round harmony pattern.

In the present study, I provide evidence that the same asymmetries observed in typologies, which are argued to derive from phonetic factors, occur as asymmetries in learning outcomes. Using an artificial grammar learning paradigm for vowel harmony (Finley & Badecker, 2008, 2009a, 2009b; Koo & Cole, 2006; Moreton, 2008; Pycha, Nowak, Shin, & Shosted, 2003), I tested the hypothesis that learners show a bias towards mid vowels as round harmony source triggers. If a learning difference exists between mid and round harmony source triggers, this evidence would suggest that phonetic factors play a role in learning round harmony patterns.

Previous artificial grammar learning studies have addressed the role of phonetic naturalness in phonological learning with mixed results (Koo & Cole, 2006; Moreton, 2008; Peperkamp, Skoruppa, & Dupoux, 2006; Pycha, et al., 2003; Schane, Tranel, & Lane, 1974; Seidl & Buckley, 2005; Wilson, 2006). For example, infants in Seidl and Buckley's (2005) study did not show any benefit for naturalness in learning. Infants in both the natural and the unnatural conditions showed a novelty preference for items that differed from the pattern they were exposed to, and there were no significant differences between conditions. Peperkamp and Dupoux (2007) found a similar result in adult learners. In a phonological rule learning task, participants showed learning for both phonetically grounded and phonetically unnatural patterns. These results support a view that learners do not share biases towards phonetically natural patterns, as the unnatural patterns were learnable in all cases. However, there are two possible alternative explanations for these results. First, it is possible that the participants learned the natural pattern faster than the unnatural pattern, but because learners were only tested at a stage when both patterns were learned, the differences were not detectable at the time of testing. Schane et al. (1974) found that when learners were tested at two stages of exposure, learners of a natural rule performed better at the first stage of exposure but not the second. This suggests that biases may appear early in training, but disappear after a sufficient amount of exposure. Second, it is possible that some phonological patterns may be more subject to naturalness constraints than others. For example, Pycha et al. (2003) found no learning differences between a harmony pattern and (a less natural) disharmony pattern. However, when learners were given an arbitrary agreement pattern (in which there was no

phonetic feature determining the relationship between the harmonic suffix and the triggering vowel), learners failed to perform above chance on test items, suggesting that different levels of naturalness may have varying effects on learnability.

Pycha et al.'s (2003) results support the notion that more complex phonological patterns may be more difficult to learn than simple phonological patterns. A pattern based on natural classes can be described with fewer rules and constraints than an arbitrary pattern. Along these lines, Peperkamp et al. (2006) found that adult participants are more likely to learn a phonotactic pattern that makes use of natural classes than an arbitrary pattern. In addition, Moreton (2008) found superior learning of a cross-linguistically common vowel height dependency compared to a cross-linguistically rare dependency between vowel height and consonant voicing This difference can be explained in terms of complexity; vowel height patterns only require a representation of vowel features, while consonant-vowel interactions involve representations of both vowel and consonant features.

Previous harmony learning experiments found some evidence for phonetic naturalness in learning. Wilson (2003) found that learners of a natural nasal harmony pattern performed better at testing than learners of an arbitrary consonant harmony pattern. Finley and Badecker (2009a) trained participants on a vowel harmony pattern in which stems triggered alternations between suffixes, alternating between [-mi] and [-mu]. Stems containing front vowels triggered [-mi] while stems containing back vowels triggered [-mu]. All participants were presented with four vowels from a six-vowel inventory. One set of participants was exposed to mid vowels (/e, o/) and high vowels (/i, u/), while another set of participants was exposed to low vowels (/ae, a/) and high vowels (/i, u/). Following exposure, participants were tested on their ability to distinguish harmonic words from disharmonic words. Test items included words from the training set, words not contained in the training set with the same vowels and consonants that appeared in the training set (New Items), and items that contained vowels not heard at training (either low vowels or mid vowels). English lacks contrasts for rounding in low vowels in both the experimental material, and in many dialects of English, (including the speaker who produced the stimuli). Thus, participants did not extend the round harmony pattern to low vowels, but were able to generalize to mid vowels. Further, Finley and Badecker (submitted) demonstrated that participants were more likely to learn a vowel height harmony pattern when the harmony undergoers were phonetically natural (and cross-linguistically attested) front vowels.

The present experiment, based on Finley and Badecker (2009a), provides evidence that learners are sensitive to the distinction between mid and high vowels as source triggers for round harmony. We exposed learners to a round harmony language with either high vowel source triggers [i, u] or mid vowel source triggers [e, o]. Participants exposed to the language with mid vowel source triggers successfully generalized the harmony pattern to novel items, but participants exposed to the high vowel source triggers did not.

Method

Participants

All participants were adult native speakers of English with no knowledge of a vowel harmony language. Sixty-eight Johns Hopkins University undergraduate students and affiliates participated either for extra course credit or for $7. The final analysis included 21 participants in each of three training conditions (Control, Mid Vowel Trigger, and High Vowel Trigger). One participant was excluded due to experimenter error. Four participants were excluded based on their performance in the perceptual task described below.

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).

The exposure phase in the critical conditions (Mid Vowel Trigger and High Vowel Trigger) was designed to be an analogue of stem-controlled vowel harmony systems, in which the vowel quality of a stem determines the vowel quality of a prefix or suffix allomorph (e.g., [-un] vs. [-in] in Turkish). Participants heard harmonic 'stems' immediately followed by a harmonic 'suffixed' form (e.g., [bidi] followed by a form containing the harmonic suffix [-ge], [bidige]). While the experimental design is discussed in terms of morphology, participants were not given information about the morphological status of words. The terms 'stem' and 'suffix' are used in the present discussion to distinguish between the material that triggers harmony (the stem) and material that alternates due to harmony (the suffix). It is not necessary that participants analyzed the stimuli in such ways to learn the harmony pattern (Finley, submitted).

There were 24 stem-suffix pairs in each critical condition. Stems were of the form CVCV[3], and contained two identical vowels. All stem vowels in the Mid Vowel Trigger Condition were mid ([e, o], as in /bede/ and /gobo/) and all stem vowels in the High Vowel Trigger Condition were high ([i, u], as in /bidi/ and /gubu/).[4] There were two different suffixes, each with two allomorphs, creating four allomorphs that were identical in both critical conditions (/mi/ vs. /mu/ and /ge/ vs. /go/). All suffixed items therefore contained three vowels. The Mid Vowel Trigger condition contained the vowel patterns /ee-i/, /ee-e/, /oo-u/, and /oo-o/, and the high vowel Trigger condition contained the vowel patterns /uu-u/, /ii-i/, /uu-o/ and /ii-e/.

The items were balanced such that 12 of the stem+suffix pairs contained only front vowels and 12 of the stem+suffix pairs contained only back vowels. Suffixes were balanced such that 12 of the items contained a high suffix vowel and 12 of the items contained a mid suffix vowel; each particular suffix allomorph appeared in 25% of the suffixed items. This yielded four evenly distributed types of stem+suffix pairs in each condition (with six tokens of each), where half of the stems were paired with a suffix that shared the same vowel height (e.g., [podo, podogo]) and half of the stems were paired with a suffix of a different vowel height (e.g., [bidi, bidige]).Stem consonants were drawn from the set: [b, d, g, p, t, k, m, n] in all conditions. The basic design is pictured in Table 1, where it can be seen that both the High Vowel and the Mid Vowel Trigger Conditions have both mid and high suffix vowels, but vary by the height of the stem vowels.

A Control condition was created to assess biases in learners outside of the harmonic training data, as well as potential anomalies in the stimuli that may inadvertently lead to above-chance performance. The purpose of the Control condition was to assess the role of training in the responses of the critical conditions. If learners respond based on pre-existing biases towards harmonic/disharmonic sequences, unrelated acoustic differences in the harmonic and disharmonic test items, or based on the exposure to the stem forms (which contained no harmonic information), the results of the critical condition should be identical to those of the Control condition. The Control condition therefore serves to ensure that any results are due

---

[3]While it would have been possible to create mono-syllabic stimuli (e.g., of the form CV or CVC), the present experiment used bi-syllabic stimuli in order to minimize the number of items that overlapped with English words.
[4]All items followed the general round harmony pattern, meaning that the Mid Vowel Trigger condition did not contain items like [piki-gu] in which high vowels fail to trigger harmony. Rather, the Mid Vowel Trigger condition only contained mid vowels in the stem.

to learning and not unforeseen, unrelated factors. Providing participants with some training makes it possible to give the same instructions (training and test), as the trained condition (as opposed to a 'no training' condition[5]).

Participants in the Control condition heard 24 harmonic unsuffixed (CVCV) stems and 24 disharmonic unsuffixed (CVCV) stems. Half of the participants heard the same harmonic stems as participants in the Mid Vowel Trigger condition and the other half heard the harmonic stems from the High Vowel Trigger condition. All participants heard the same disharmonic stems. Participants received identical test items in each of the critical conditions, described below. Those who heard harmonic stems from the High Vowel Trigger condition received test items from the High Vowel Trigger condition (and likewise for the Mid Vowel Trigger condition). While all test items were technically 'new' (as no participant in the Control condition heard suffixed items), all test items were assigned to the appropriate condition based on the corresponding critical condition.

All participants received a two-alternative forced-choice test immediately following exposure. Each test item contained two suffixed items, one harmonic and one disharmonic (e.g., [*bidi**mu** vs. bidi**mi**]). Test items included items from the training set (Old) and items not included in the training set (New), depicted in Table 2[6]. The New test items were designed to test whether participants learned an abstract pattern, or had simply memorized the forms heard in training. If learners are sensitive to the phonetic grounding of source triggers for round harmony, participants in the Mid Vowel Trigger condition should show greater harmonic responses compared to both the High Vowel Trigger condition and the Control condition, particularly for New test items, as New test items directly test for learning of an abstract pattern.

The order of presentation of harmonic and disharmonic options in the test items was counterbalanced such that the correct (harmonic) item was presented first for half of the items, and presented second in the other half of items. In addition, the set of correct items contained a round vowel in half of the test items, and an unround vowel the other half of the test items. All suffixes were familiar to the participants and were divided evenly such that half of the items ended in [-ge]/[-go] and half ended in [-mi]/[-mu].All participants were screened with an AXB perceptual task, focusing on perception of English vowel height and rounding features. Participants judged whether the first or the third syllable contained a vowel identical to the vowel in the second syllable. For example, if participants heard [be], [be] and [bo], the correct response would be the first syllable ([be]). All items were monosyllabic, and contained the vowels /i, e, u, o/. Those (four) participants scoring less than 75 percent were removed from the study, with the logic that all participants should have general competence for perceiving English vowels in order for their learning data to be meaningful.

The general design of the study has many similarities to Finley and Badecker (2009a). First, both studies presented materials in the exposure phase as stem-suffix pairs. Second, both studies made use of two different training conditions (Mid Vowel Trigger) and (High Vowel Trigger). Third, all items in the critical conditions showed agreement in terms of round features. However, in Finley and Badecker (2009a), participants heard stems from two different vowel heights. In the present study, all stem vowels shared the same vowel height. All vowels in the Mid Vowel Trigger Condition, were mid (e.g., [bede]), and all stem

---

[5]Previous research (e.g., Finley and Badecker (2009a)) found no differences between 'no-training' control conditions and stem-only control conditions.
[6]To replicate Finley and Badecker (2009a), we included New Vowel Items in the test set. However, these items are not included in the analysis because the relevant comparison was between Old and New items; participants were at chance for all New Vowel items.

vowels in the High Vowel Trigger condition were high (e.g., [piki]). Another difference between Finley and Badecker (2009a) is that the present experiment made use of both mid and high vowel suffix alternations for all conditions (e.g., [-mi]/[- mu] and [-ge]/[-go]), while Finley and Badecker used a single suffix alternation (e.g., [-mi]/[-mu]). Finally, the hypotheses of the present paper differ significantly from Finley and Badecker. Finley and Badecker (2009a) tested whether learners could generalize to novel stem vowels, while the present study tests the learnability of harmony with non-ideal source triggers.

Materials

The naturally produced stimuli were recorded in a sound proof booth at 22,000kHz from an adult male speaker of American English with basic phonetic training (he had completed a graduate-level phonetics course). While the speaker had no knowledge of the specifics of the experimental design, he was aware that the items would be used in an artificial language learning task. All stimuli were phonetically transcribed, and presented to the speaker in written format. The speaker was instructed to produce all vowels as clearly and accurately as possible, even in unstressed positions. The speaker was told to produce all sounds equally clearly. A rater judged each word, and any word that did not have clear vowel sounds was re-recorded for clarity. Stress was produced on the first (stem) syllable in all forms.

Suffixed items were recorded as CVCVgə (as opposed to CVCV-gi or CVCV-gu) The final schwa ensured that the speaker produced minimal coarticulation on the final stem vowel. It also ensured that the stem portion of the word would be identical for both test items in the two-alternative forced-choice test. Suffix and stem allomorphs were therefore created with a single token. For example, to create the test item containing [bidigi] and [bidigu], the speaker produced three different words: [bidigə], [bəbəgi] and [bəbəgu]. The harmonic [bidigi] was taken by cross-splicing the [bidi] portion of [bidigə] with the [gi] portion of [bəbəgi]; the disharmonic [bidigu] was created by cross-splicing the [bidi] portion of [bidigə] with the [gu] portion of [bəbəgu]. This ensured that the only difference between the two test items was the final vowel, which was either harmonic or disharmonic. All sound editing was performed with Praat (Boersma & Weenink, 2005). Intensity was scaled to 70dB for all words. We measured pitch and duration to ensure there were no significant differences between mid and high vowels. Pitch was not significantly different across vowels (average of 107 Hz for high vowels and 105 Hz for mid vowels, t<1.) Duration was not significant across vowels (average of 146 ms for high vowels and 145 ms for mid vowels, t< 1).

Procedure

Participants were randomly assigned to one of three conditions: a Control condition, a High Vowel Trigger condition and Mid Vowel Trigger condition. 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, but that they need not memorize the forms. Participants listened to the 24 training item stem-suffix pairs repeated five times in a different random order for each cycle.

Following training, participants received the 24-item forced-choice test. Participants were told that they would hear two words, one from the language they just heard, the other not from the language; their job was to select the word from the language. They were told to press the 'a' key if they thought the first word was from the language, and to press the 'l' key if they thought the second word was from the language. They were told to respond as quickly and accurately as possible, but to wait until they heard both possibilities before

responding. Participants were then given the AXB perception task for English vowels described above. The experiment took approximately 15 minutes to complete.

Results

Proportion of harmonic responses are given in Figure 1. The data was analyzed using a 3×2 ANOVA with Training (Mid Vowel Trigger, High Vowel Trigger and Control) as the between-subject factor and Test Item (Old and New) as the within-subject factor. There was a main effect of condition, $F(2, 60) = 6.11$, $p < 0.01$, with no effect of Test Item, $F < 1$, and no interaction, $F < 1$. Pair-wise comparisons between each condition revealed a significant effect of training for the Mid Vowel Trigger condition compared to the Control condition, with means of 0.66 and 0.51, respectively ($CI± 0.082$, $p < 0.01$), but no significant difference between the High Vowel Trigger condition and the Control condition, with means of 0.58 and 0.51, respectively ($CI± 0.093$, $p = 0.11$). There was a marginally significant difference between the Mid Vowel Trigger and the High Vowel Trigger conditions, with means of 0.66 and 0.58, respectively ($CI± 0.082$, $p = 0.062$). This marginal effect was carried by the fact that there was a significant difference between the Mid and High Vowel Trigger conditions for New test items, $t(40)=2.43$, $p < 0.05$, (corrected for multiple comparisons), but not Old test items, $t < 1$. Because above chance performance can be achieved on Old test items simply through memory of the words heard at training, New test items are critical to establish learning of a novel phonological pattern[7].

In addition, we compared responses to New Items between each critical condition and the Control condition. There was a significant difference between the High Vowel Trigger condition and the Control condition, (with means of 0.67 and 0.47, respectively, $CI± 0.11$) $t(40) = 3.71$, $p < 0.001$, but there was no difference between the Mid Vowel Trigger condition and the Control condition, (with means of 0.55 and 0.54, respectively, $CI± 0.10$) $t(40)=1.39$, $p = 0.17$. These results suggest that participants in the Mid Vowel Trigger Condition were more likely to form a general harmony pattern than participants in the High Vowel Trigger condition. Because it is possible to perform above chance on Old items without learning a general harmony pattern (as Old items were heard in training), it is not surprising that there was no difference between Mid and High vowel trigger conditions for Old items.

It is important to note that the failure of the High Vowel Trigger condition to show significant results compared to the Control condition was not carried by one or two aberrant participants. Twelve of 21 participants in the High Vowel Trigger Condition had means lower than 58% for New Items, while only 6 of 21 participants in the Mid Vowel Trigger Condition had means lower than 58% for New Items.

AXB test results were at ceiling: those scoring above threshold scored an average of 93.79% correct. The four participants who scored below 75% on the AXB task scored an average of 48.68% correct. We did not score these participants on their learning because if a native English speaker fails this very easy task, they are either not attending or did not understand the task. Therefore, their overall data cannot be interpreted.

---

[7]These results were the same when we compared the Mid Vowel Condition to its matched Control condition, $F(1,30) = 10.09$, $p = 0.003$, with no interaction and no effect of test item, $F < 1$, or compared the High Vowel Trigger Condition to its matched control condition (there was no effect of Training, $F < 1$, Test Item, $F < 1$, and there was no interaction $F < 1$).

DISCUSSION

Understanding biases in language learning can increase understanding of general language learning mechanisms. The results from the present experiment suggest that the ability to learn a round vowel harmony pattern is improved if the pattern contains an ideal round harmony source trigger. Participants were able to learn the round vowel harmony pattern when exposed to mid vowel source triggers only, but participants failed to learn the pattern when exposed to high vowel source triggers only. These results support the view that learners are sensitive to phonetic factors that drive the typology of vowel harmony.

The evidence suggests that the role (if any) of the frequency of mid versus high vowels is minimal. According to frequency counts (Kessler & Treiman, 1997), there does not appear to be any major differences between mid and high vowels for English (the native language of the participants in this study). Nor does there appear to be a difference between front and back vowels (front vowels appear to be slightly more frequent). Thus, it is unclear what (if any) effect phoneme frequency has on the present results.

Harmonic and disharmonic sequences between mid and high vowels (on the round dimension) were compared using a searchable English pronunciation dictionary (Szigetvári, 2009). I counted pairs of vowels, comparing harmonic sequences beginning with a high vowel to harmonic sequences beginning with a mid vowel. There were relatively equal numbers of words containing harmonic sequences with mid vowels (4,355) compared to high vowels (4,901). There were also relatively equal numbers of disharmonic pairs beginning with mid vowels (2,798) compared with sequences beginning with high vowels (2,867). Overall there were greater harmonic sequences than disharmonic sequences. The greater number of harmonic sequences is carried mainly by harmonic sequences with the reduced vowel /ə/, which was not used in this study, and is highly frequent in English. When only vowel sequences used in the study were considered, the differences were reduced, but still slightly favored harmonic sequences (164 to 145 for mid vowels and 142 to 110 for high vowels). However, these frequency counts must be interpreted with caution, as they are not weighted for token frequency and the corpus used transcribed many unstressed vowels as /ɪ/ or /ə/. Further, if participants used only the frequency counts in English (as opposed to the training) they were exposed to, we would not expect differences between the conditions. This is especially true of the Control condition, whose participants were at chance, suggesting no preference for harmonic or disharmonic items. These counts demonstrate that the patterns for mid and high vowels are roughly equivalent, and the small differences that are found do not provide an alternative explanation to the pattern of data collected in the present study.

Confusability may have played a role in producing the present results. It is possible that listeners in the High Vowel Trigger Condition confused /ii-i/ training items as disharmonic /ii-u/ items. This confusion could have occurred based on two properties of English. First, English /i/ exerts a relatively strong coarticulatory influence on surrounding vowels (Cole et al., 2010). Second, because the English vowel /u/ is often fronted in many dialects of English (Labov, Ash, & Boberg, 2006), speakers may have inferred that /ii-i/ items were actually /ii-u/ items as a consequence of compensation for coarticulation. If participants in the High Vowel Trigger Condition misheard training items as /ii-u/ rather than harmonic /ii-i/, it would impede learning. While testing for this possibility would be extremely difficult, this alternative explanation is in line with the general hypothesis presented in this paper that fine-grained phonetic details can affect learning. While the exact phonetic precursors for learning biases may be in question, it is clear that phonetic effects can bias learners towards some patterns over others, and that these biases may reflect the distribution of sound patterns across languages of the world.

One question that remains from the present results is why the AXB perception test showed no significant differences between mid and high vowels. It is possible that a more difficult perceptual task may have revealed differences. However, given that rounding co-occurs with the back feature in English, speakers are unlikely to misperceive round/back counterparts such as [e] and [o] because they have multiple feature cues, which is independent of the perceptual benefit from perceiving a vowel as round in a harmonic context. Further, the differences between mid and high vowels as source triggers were manifested via learning. This was confirmed by the fact that the Control condition showed no significant differences between harmonic responses for /i/ compared to /u/. This suggests that perception of round versus unround vowels in isolation may not be an accurate predictor of the phonetic precursor to round harmony source triggers.

The present results provide evidence for a bias against poor harmony triggers for round vowel harmony. The question remains as to where this bias comes from, and how it emerges. There is evidence that phonetic biases are shaped through experience with language (Gerken & Bollt, 2008; Kuhl, 2001; Kuhl, Williams, Lacerda, Steven, & Lindblom, 1992). For example, Gerken and Bollt (2008) showed that younger infants (7.5 months) are more likely to show generalizations based on unnatural stress patterns than older infants (9 months), who only made generalizations based on natural stress patterns. Therefore, adults whose phonetic knowledge is fully formed may provide the most accurate account of human learning biases. If learning biases are shaped through experience, the strongest biases should be found in adult learners, who have the greatest amount of exposure to language. Studying learning biases in adults may provide some insight into first language learning. The biases found in the present experiment are based on the phonetic underpinnings of vowel harmony, and should therefore apply to learners of various ages. A question for future research is when these biases emerge. Infants' experiences with language provide insight into the way languages are structured, making it more likely for an older infant to show naturalness preferences than younger infants. It is possible that as children learn about the vowel system of their languages, they become aware of the perceptual and articulatory mechanisms that bias the sound patterns of language, in addition to their native language. Future work will address the address the role of naturalness in learning over time.

CONCLUSION

The present paper provided evidence that learners are sensitive to phonetic biases in learning phonological patterns for round vowel harmony. While both sets of learners were exposed to nearly identical harmony patterns, only learners who were exposed to the phonetically motivated round harmony pattern were successful in learning. Because I compared two harmony languages that shared identical structural properties (only differing by the height of the stem vowels), it is clear that the effects of training found in this study are not due to complexity. Learning in an artificial grammar setting is rapid and robust, but this learning is sensitive to the phonetic grounding of the phonological pattern.

Acknowledgments

References

Anderson S. Why phonology isn't "natural". Linguistic Inquiry. 1981; 12:493–547.

Blevins, J. Evolutionary phonology: The emergence of sound patterns. Cambridge University Press; Cambridge: 2004.

Boersma, P.; Weenink. Praat: Doing phonetics by computer. 2005.

Clements, GN.; Sezer, E. Vowel and consonant disharmony in Turkish. In: Smith, v. d. H. a., editor. The structure of Phonological Representations. Vol. Vol. II. Foris; Dordrecht: 1982. p. 213-255.

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.

Evans N, Levinson SC. The myth of language universals: Language diversity and its importance for cognitive science. Behavioral and Brain Sciences. 2009; 32(5):429–492. doi: 10.1017/ S0140525X0999094X. [PubMed: 19857320]

Finley, S. Generalization to novel talkers in artificial grammar learning. (submitted)

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. Learning biases for vowel height harmony. (submitted)

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

Guest DJ, Dell GS, Cole J. Violable constraints in language production: Testing the transitivity assumption of Optimality Theory. Journal of Memory and Language. 2000; 42:272–299.

Kaun, A. The typology of rounding harmony: An optimality theoretic approach. UCLA; Los Angeles: 1995.

Kaun, A. The typology of rounding harmony. In: Hayes, B.; Steriade, D., editors. Phonetics in Phonology. Cambridge University Press; Cambridge: 2004. p. 87-116.

Kessler B, Treiman. Syllable structure and the distribution of phonemes in English syllables. Journal of Memory and Language. 1997; 37:295–311.

Koo, H.; Cole, J. On learnability and naturalness as constraints on phonological grammar. In: Botinis, A., editor. Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics. Athens: 2006. p. 174-177.

Korn D. Types of labial vowel harmony in the Turkic languages. Anthropological Lingusitics. 1969; 11:98–106.

Kuhl, PK. Language, mind and brain: Experience alters perception. In: Nelson, C.; Luciana, M., editors. Handbook of developmental cognitive neuroscience. MIT Press; Cambridge: 2001. p. 99-115.

Kuhl PK, Williams KA, Lacerda F, Steven K, Lindblom B. Linguistic experience alters phonetic perceptoin in infants by 6 months of age. Science. 1992; 255:606–698. [PubMed: 1736364]

Labov, W.; Ash, S.; Boberg, C. The atlas of North American English: phonetics, phonology, and sound change. Mouton de Gruyter; Berlin: 2006.

Linker, W. Articulatory and acoustic correlates of labial activity in vowels: A cross-linguistic study. Vol. 56. UCLA; Los Angeles: 1982. UCLA Working Papers in Phonetics

Moreton E. Analytic bias and phonological typology. Phonology. 2008; 25:83–127.

Nevins A. On formal universals in phonology. Behavioral and Brain Sciences. 2009; 32:461–462.

Peperkamp, S.; Dupoux, E. Learning the mapping from surface to underlying representations in artificial language. In: Cole, J.; Hualde, JI., editors. Laboratory Phonology. Vol. 9. 2007. p. 315-338.

Peperkamp, S.; Skoruppa, K.; Dupoux, E. Boston University Conference on Language Development. Vol. Vol. 30. Cascadilla Press; 2006. The role of phonetic naturalness in phonological rule acquisition; p. 464-475.

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.

Schane S, Tranel B, Lane H. On the psychological reality of a natural rule of syllable structure. Cognition. 1974; 3:351–358.

Seidl A, Buckley E. On the learning of arbitrary phonological rules. Language Learning and Development. 2005; 1:289–316.

Suomi K. Palatal vowel harmony: A perceptually motivated phenomenon? Nordic Journal of Linguistics. 1983; 6:1–35.

Szigetvári, P. Searchable English pronunciation dictionary. 2009. http://seas3.elte.hu/epd.html

Terbeek D. A cross-language multidimensional scaling study of vowel perception. UCLA Working Papers in Linguistics. 1977; 37:1–271.

van der Hulst, H.; van de Weijer, J. Vowel harmony. In: Goldsmith, J., editor. The handbook of phonological theory. Blackwell; Oxford: 1995. p. 495-531.

Walker R. Weak triggers in vowel harmony. Natural Language & Linguistic Theory. 2005; 23:917–989.

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]
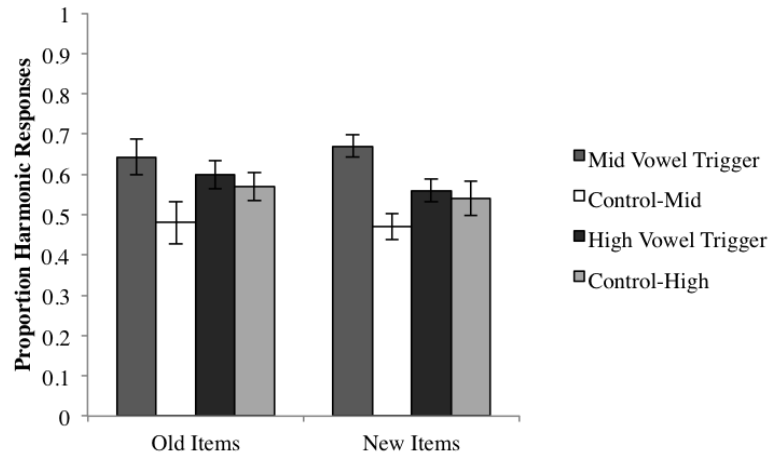
Figure 1.

Table 1

Training Design: (Stem Vowels-Suffix Vowel)

High Vowel Trigger		Mid Vowel Trigger	
Front Suffix Vowel	**Back Suffix Vowel**	**Front Suffix Vowel**	**Back Suffix Vowel**
ii-e	uu-o	ee-e	oo-o
ii-i	uu-u	ee-i	oo-u

Table 2

Test Stimuli

Test condition	High Vowel Trigger	Mid Vowel Trigger
Old	bidimi vs. *bidimu	*gomomi vs. gomomu
	nupugo vs. *nupuge	*bodoge vs. bodogo
New	pidimi vs. *pidimu	mepemi vs. *mepemu
	*nukuge vs. nukugo	*gotoge vs. gotogo