

Generalization to unfamiliar talkers in artificial language learning

Sara Finley

Published online: 28 February 2013
© Psychonomic Society, Inc. 2013

Abstract While there is evidence that talker-specific details are encoded in the phonetics of the lexicon (Kraljic, Samuel, & Brennan, *Psychological Science* 19(4):332–228, 2008; Logan, Lively, & Pisoni, *Journal of the Acoustical Society of America*, 89(2):874–886, 1991) and in sentence processing (Nygaard & Pisoni, *Perception & Psychophysics*, 60(3):355–376, 1998), it is unclear whether categorical linguistic patterns are also represented in terms of talker-specific details. The present study provides evidence that adult learners form talker-independent representations for productive linguistic patterns. Participants were able to generalize a novel linguistic pattern to unfamiliar talkers. Learners were exposed to spoken words that conformed to a pattern in which vowels of a word agreed in place of articulation, referred to as *vowel harmony*. All items were presented in the voice of one single talker. Participants were tested on items that included both the familiar talker and an unfamiliar talker. Participants generalized the pattern to novel talkers when the talkers spoke with a familiar accent (Experiment 1), as well as with an unfamiliar accent (Experiment 2). Learners showed a small advantage for talker familiarity when the words were familiar, but not when the words were novel. These results are consistent with a theory of language processing in which the lexicon stores fine-grained, talker-specific phonetic details, but productive linguistic processes are subject to abstract, talker-independent representations.

Keywords Artificial grammar learning · Phonology · Psycholinguistics

Language use involves knowledge of highly specific phonetic details, as well the ability to generalize to novel

situations, raising the issue of the extent to which knowledge of language relies on abstract rules versus fine-grained details. Knowledge of a specific language is general; a language user can understand almost any speaker of the language, despite the fact that every speaker has individual, idiosyncratic characteristics. Thus, the language user must be able to distinguish between speech characteristics that are idiosyncratic to the talker and speech characteristics that are shared across the language. Studying how speakers deal with unfamiliar talkers in language learning tasks can help to uncover which aspects of language processing make use of talker-specific information and which aspects of language take place at a talker-independent level of representation.

Previous research exploring the role of talker-specific effects of language processing has focused on lexical access and phonetic patterns, with little discussion of productive, categorical linguistic patterns. This article focuses on productive phonological patterns: systematic changes in the sounds that make up a word. For example, vowel harmony is a phonological pattern that can be found in several of the world's languages but is not found in English. With some exceptions, Hungarian shows alternations in suffix vowels depending on the quality of the stem vowels.¹ For example, the singular (dative) suffix alternates between [-nek] and [-nak], depending on the quality of the stem vowel. When the stem contains vowels pronounced in the back of the oral cavity, such as /a/ and /o/, [-nak] appears (e.g., [hajo-nak] 'ship'). When the stem contains vowels produced in the front of the oral cavity, such as /i/ and /e/, [-nek] appears (e.g., [öleles-nek] 'embracement').

In Hungarian, the formation of morphologically complex words is dependent on vowel harmony, demonstrating the interaction between phonological patterns and the lexicon. This interaction has led some researchers to propose the possibility of reducing the study of productive phonological patterns to tendencies over the lexicon (Port & Leary, 2005).

S. Finley (✉)
Department of Psychology, Waldorf College, 106 S 6th Street,
Forest City, IA 50436, USA
e-mail: sara.finley@waldorf.edu

¹ In the word *dogs*, *dog* is the stem and *-s* is the suffix.

Exemplar models of cognition (Goldinger, 1996, 1998; Nosofsky, 1988) serve as the basis for many of these proposals (Connine & Pinnow, 2006; Johnson, 1997; Palmeri, Goldinger, & Pisoni, 1993; Pierrehumbert, 2001; Wedel, 2006). Exemplar models of language are supported by the robust finding that talker familiarity serves as an aid to lexical access (Connine & Pinnow, 2006; Creel, Aslin, & Tanenhaus, 2008; Creel & Tumlin, 2009; Goldinger, 1998; Palmeri et al., 1993; Pisoni, 1997). In these studies, listeners were faster and more accurate when recalling words spoken by a familiar talker, rather than by an unfamiliar talker. These results suggest that talker-specific information is encoded in the lexicon. In addition, each token of a word is stored in memory along with the fine-grained phonetic characteristics of that token.

Because lexical entries encode talker-specific information, there is reason to believe that phonological patterns may also make use of talker-specific information (Pierrehumbert, 2001). While no studies have focused on categorical phonological patterns like vowel harmony, research has tested learners' ability to generalize to novel talkers when learning a novel phonetic contrast, such as the difference between /l/ and /r/ that is present in English but not Japanese. Japanese learners of English were able to extend the novel phonetic category to novel talkers only when participants were trained on stimuli that included multiple talkers (Lively, Pisoni, & Logan, 1992; Logan, Lively, & Pisoni, 1991; Magnuson et al., 1995). This suggests that phonetic contrasts may be formed via talker-specific representations. However, categorical phonological patterns like vowel harmony may be represented differently than phonetic contrasts that tend to make greater use of fine-grained phonetic details. It is therefore unclear whether categorical phonological patterns like vowel harmony will show the same talker-specific effects.

There is evidence that talker-specific knowledge is used in abstract, sentence-level processing (Nygaard & Pisoni, 1998). Participants were trained on talker identities by listening to several talkers produce full sentences. Participants showed better recognition for individual words in test sentences that were spoken by familiar talkers, demonstrating that talker-specific information is used in high-level language processing. These results support the view that linguistic processes, including phonological patterns, may be subject to talker-specific effects. Because Nygaard and Pisoni only tested word recognition, it is unclear whether talker-specific processing extends to processing novel stimuli from categorical phonological patterns like vowel harmony, when participants are trained on a single phonological pattern spoken by one individual talker.

In addition, phonetic differences between categories of sounds may make some sounds more applicable to generalization to novel talkers. In a perceptual learning task, Kraljic and Samuel (2006) showed that participants can extend a novel category boundary to unfamiliar talkers but that this generalization is constrained by the specific sounds

involved in the contrast (Kraljic & Samuel, 2005, 2007), as well as the particular behavior of the talker (Kraljic, Samuel, & Brennan, 2008). Speakers were less likely to extend the novel accent to an unfamiliar talker if the familiar talker spoke with a pen in the mouth (Kraljic et al., 2008).

This article uses an artificial grammar learning paradigm to study the effects of talker familiarity on processing novel phonological patterns. Previous artificial language learning experiments have shown that adults can learn a novel vowel harmony pattern after brief exposure (Finley & Badecker, 2008, 2009a, 2009b, 2010; Koo & Cole, 2006; Moreton, 2008; Pycha, Nowak, Shin, & Shosted, 2003; Skoruppa & Peperkamp, 2011). In addition, learners of novel phonological patterns show robust generalization to novel items. For example, participants in a study by Finley and Badecker (2009a) heard a novel back/round vowel harmony pattern that contained four vowels from a six-vowel inventory. Following exposure, participants were given a two-alternative forced choice task in which participants chose between a harmonic form and a disharmonic form. For example, participants chose between harmonic [bodumu] and disharmonic *[bodumi].² Test items were divided into three categories: *old items*, which were words that appeared in training; *new items*, which were words that did not appear in training but contained the same vowels and consonants as the training set; and *new vowel items*, which were words that contained vowels not present in the training set. Participants extended the vowel harmony pattern to novel vowels, suggesting that novel phonological processes are learned in terms of abstract representations.

Previous artificial language learning experiments, specifically those exploring phonological patterns, made use of a single talker during exposure. These studies did not test for generalization to novel talkers. This leaves open the possibility that learners infer talker-specific representations rather than language-specific representations. Some insight to how learners respond to novel talkers in an artificial language learning task may be gained through examination of artificial grammar learning experiments that explored the role of transfer across modalities, such as from a spoken pattern to an analogous written pattern (Dienes & Altmann, 1997). These studies found robust transfer across modalities, but also found a transfer deficit; correct responses decreased in the novel modality. While transfer across modalities is different from transfer across talkers, understanding the level of representation for which items are stored tests the limits of human generalization. These findings are useful in creating theories of the levels of representation for both linguistic and nonlinguistic pattern learning. In addition, the existence of a *transfer deficit* in generalization across modalities suggests that learners will show deficits across talkers as well (Dienes & Altmann, 1997).

² The “*” indicates an ungrammatical, disharmonic item.

The present study tests for the existence of talker-independent representations in novel phonological patterns. Evidence for talker-independent representations will be found if participants are able to extend the novel harmony pattern from a familiar talker to an unfamiliar talker. This extension will be measured in two ways. First, if correct responses to the unfamiliar talker exceed correct responses in the control condition, it suggests that participants have a representation of the vowel harmony pattern that goes beyond familiar talker. Second, transfer deficits will be assessed by comparing responses to the unfamiliar talker with corresponding responses to the familiar talker. Statistically significant differences between familiar and unfamiliar talkers provide evidence for talker-specific details within the representation of the harmony pattern. A division between old items (words heard in training) and new items (novel words not heard in training) allows for comparisons to be made with respect to lexical familiarity. These different comparisons yield several possible outcomes and interpretations. The four most probable outcomes are listed in Table 1.

If participants in the experimental condition fail to extend the pattern to unfamiliar talkers, as compared with the control condition, it suggests that the harmony pattern was learned using talker-specific representations. If there is a failure to extend the harmony pattern to unfamiliar talkers, one expects that there will also be significant transfer deficits for both old and new items. Because failure to extend to unfamiliar talkers has been shown in previous studies, (Kraljic & Samuel, 2007), it is possible that learners in the present study will also fail to extend the harmony pattern to unfamiliar talkers.

If, in addition to storing information about the familiar talker, speakers have access to a general, talker-independent representation of the pattern, learners may show a significant generalization to unfamiliar talkers, as compared with the control condition, but with transfer deficits. Under this pattern of results, there are two possible outcomes: transfer deficits for both old and new items or transfer deficits for old items only. If transfer deficits occur for all items, it suggests that learners make use of talker-specific details to learn the harmony pattern, but when exposed to the same pattern spoken in an unfamiliar voice, the learner must extend the stored representations that contain the familiar

talker to the unfamiliar talker, resulting in a transfer deficit for all items. However, it is also possible that transfer deficits will occur only for old items. Much of the research supporting talker-specific representations has focused on lexical access. There is evidence that words are easier to access if they are spoken by a familiar talker (Nygaard, Sommers, & Pisoni, 1994; Palmeri et al., 1993). If the general phonological pattern is learned under talker-independent mechanisms but the lexicon incorporates talker-specific information, one should expect that participants will correctly respond to items spoken with an unfamiliar talker (as compared to a control condition) but will show transfer deficits for old items only. New items will not show transfer deficits, because the learner has no prior representation of these words in the lexicon.

The final possible outcome assumes that phonological patterns are stored without any talker-specific information. In this case, participants will extend the harmony pattern to unfamiliar talkers without any transfer deficit. This pattern of results would support the strongest version of talker independence. In order for this outcome to occur, learners must show the same pattern of results for the familiar talker and the unfamiliar talker, even for familiar (old) test items. This possible pattern of results supports a view in which both the lexicon and the phonological pattern make use of talker-independent representations.

The experiments in the present article explore the role of talker independence in learning a novel phonological pattern. The ability to extend a novel vowel harmony pattern to an unfamiliar talker will help to shed light on the nature of learning and representations of phonological knowledge.

Experiment 1

Method

Participants

All participants were adult native speakers of English with no knowledge of a vowel harmony language. Seventy-two

Table 1 Possible results and interpretations

Representations for phonological patterns	Extend pattern to unfamiliar talkers (compared to control)	Transfer deficit	
		Old items	New items
Talker specific	No	Yes	Yes
Talker-specific representations extended to unfamiliar talkers	Yes	Yes	Yes
Lexicon is talker specific but patterns are talker independent	Yes	Yes	No
Talker independent	Yes	No	No

University of Rochester undergraduate students and affiliates were paid \$10 for their participation. There were 40 participants in the critical conditions and 32 participants in the no-training control condition.

Design

An artificial grammar learning task was used to assess the ability to generalize novel phonological patterns to novel talkers. In an artificial grammar task, a novel language is created that has a specific pattern or characteristic, such as vowel harmony. Participants in the experimental conditions were exposed to a vowel harmony pattern spoken by a single talker, either male or female. The training phase was followed immediately by a two-alternative forced choice test that contained items spoken by the familiar talker, as well as a novel talker. The general design of the experiment was based on Finley and Badecker (2009a), described above. All phases of the experiment were presented using PsyScopeX (Cohen, MacWhinney, Flatt, & Provost, 1993).

Participants in the critical conditions were assigned to either the *male talker training* condition or the *female talker training* condition. Tokens in both conditions were identical, except that tokens in the male talker training condition were spoken by a male talker, while tokens in the female talker training condition were spoken by a female talker.

Training items in the critical conditions were of the form of a bisyllabic *stem* word (of the form CVCVC, where C is a consonant and V is a vowel) immediately followed by a back/round harmonic *suffixed* word (of the form CVCVC-V), but without any meanings associated with the items. For example, the stem [gemit] was followed by the harmonic [gemit-e]. Stems contained front vowels [i, e] or back vowels [o, u]. The suffix was a vowel that alternated between [-e] and [-o]. The suffix [-e] appeared when the stem vowels were front (e.g., [mekin, mekin-e]); the suffix [-o] appeared when the stem vowel was back (e.g., [poduk, poduk-o]). Examples of training stimuli can be found in Table 2; full stimuli lists are in the Appendix.

All stimuli contained the same consonant inventory, [p, b, t, d, k, g, m, n] and the same vowel inventory, [i, u, e, o].

Sixteen training items were created for each critical condition. The training stimuli were counterbalanced to contain all possible combinations of vowel sounds. Suffixed words were produced semirandomly, with the condition that all stimuli were not homophonous with an English word. The final profile of the stimuli was counterbalanced to appropriately contain equivalent use of the different vowel patterns in the stems.

Training was followed by a two-alternative forced choice test with 40 test items. Each test item contained two suffixed forms, one harmonic and one disharmonic. For example, participants chose between harmonic [gemit-e] and disharmonic *[gemit-o]. Half of the test items were presented in the female voice, while the other half of the test items contained the male voice. One group of participants heard all 40 items in an unblocked, random order, with familiar and unfamiliar talkers mixed in a random order ($n = 16$). All other participants ($n = 24$) heard test items presented in two blocks of 20 items, an all-male and an all-female block, with the order of presentation counterbalanced such that half of participants heard the male test items first.³ Each block contained 10 words that had appeared in the exposure phase and 10 words that had not appeared in the exposure phase; these items were presented in a random order. This amounted to four total test conditions: old items (familiar talker), new items (familiar talker), old items (unfamiliar talker), and new items (unfamiliar talker).

Thirty-two participants were assigned to a no-training control condition. The control condition was designed to ensure that all effects were due to learning, as opposed to an inherent response bias. In this no-training control condition, participants received the same test items as participants in the critical conditions. Half of the control participants responded to items in an unblocked, random order, while the other half responded to items in blocks. Of these participants, 8 responded to female items first, 8 participants responded to male items first, and 16 participants responded to male and female items in a random order. While the control participants had not heard any of the test items during the exposure phase and, thus, all were “new,” the test items were matched to the appropriate test condition on the basis of the critical condition.⁴

As is noted in Table 1 and the description of the table, the extent to which learning novel, categorical phonological

Table 2 Examples of training and test stimuli

Training	[netep, netepe]
	[gemit, gemite]
	[kukop, kukopo]
	[monuk, monuko]
Old items	[gemite, *gemito]
	[*kukope, kukopo]
New items	[bedite, *bedito]
	[mukobo, *mukobe]

³ Results did not indicate any difference in responses depending on blocked or random order of presentation. The overall average response rate was .78 for random and .76 for blocked.

⁴ Finley and Badecker (2009a) used a “stem” only, as well as a monosyllabic control condition. In these conditions, participants are given a nonharmonic (neither harmonic or disharmonic) pattern to learn. Responses to these control conditions were also close to chance, making it unlikely that the control condition in the present experiment was any better or worse than previous control conditions.

patterns is based in talker-independent representations should be reflected in the extent to which learners are able to extend the harmony pattern to an unfamiliar talker, both in comparison with the control condition as compared with responses to a familiar talker, as a transfer deficit.

Materials

The naturally produced stimuli were recorded in a sound-attenuated booth with a 22-kHz sampling rate from two native speakers of American English, one male and one female. Both speakers spent the majority of their childhood in the same region of the United States, upstate New York. While the speakers had no knowledge of the specifics of the experimental design, they were aware that the items would be used in an artificial language learning task. All stimuli were phonetically transcribed and presented to the speakers in written format. The speakers were instructed to produce all vowels as clearly and accurately as possible, even in unstressed positions. Stress was produced on the first syllable in all forms. Because talkers were told to speak naturally, length of utterances was not controlled for. Thus, there were differences in durations between the male and the female talkers, with the female talker being slightly longer. Male suffixed items averaged 835 ms, with a range of 760–1,000 ms. Female suffixed items averaged 1,303 ms, with a range of 881–1,401 ms. Such differences in length make the talkers even more distinguished, which, if anything, should prevent generalization to the novel talker. All items were scaled to the same intensity level. All sound files were edited in Praat (Boersma & Weenink, 2005).

Procedure

All phases of the experiments took place at a Macintosh computer with stimuli presented using over-ear headphones. Participants were allowed to adjust the volume of the headphones to a comfortable level. Participants received both written and verbal instructions. Participants in the critical conditions were told that they would be listening to words from a language they had never heard before but that they need not memorize the forms. Following the exposure phase, participants in the critical conditions 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. If they believed that the first word was from the language, they were to press the “a” key on the keyboard, and if they believed that the second word was from the language, they were to press the “l” key on the keyboard. They were told to respond as quickly and accurately as possible, but to wait until they heard both possibilities before responding. The experiment took approximately 15 min.

Participants in the control condition were only given the test items and were therefore not given any exposure items to compare with during testing. This means that the directions given to participants in the critical conditions were not appropriate for the control condition. Instead, participants were told that they would be making judgments about pairs of words. Their task was to decide which of two words they preferred, on the basis of any criteria they chose, and they were told that there was no “correct” or “incorrect” response. Participants in the control condition were given the same set of test items and responded using the computer keyboard in the same manner as participants in the critical conditions. The control experiment took approximately 5 min to complete.

Results

Proportions of same-language, harmonic responses are provided in Table 3. Results are categorized in terms of the control condition and the two critical conditions—male talker training and female talker training—and divided by old items and new items for familiar and unfamiliar talker test items. Control items have male as the default familiar talker, but statistical comparisons were made according to the appropriate gender.

A 2 (training gender) \times 2 (talker familiarity) \times 2 (lexical familiarity) mixed-design ANOVA was used to compare responses in the female talker training condition with those in the male talker training condition. This comparison ensured that responses did not differ on the basis of the gender of the talker heard during the exposure phase. There was no effect of gender, $F(1, 38) = 1.75, p = .19, \eta^2 = .044$, suggesting that participants in both training conditions were equally able to learn the harmony pattern. There was a significant effect of lexical familiarity, $F(1, 38) = 29.83, p < .0001, \eta^2 = .44$, and no interaction with gender, $F < 1$, suggesting that participants were more accurate with old items than with new items. There was a marginally significant effect of talker familiarity, $F(1, 38) = 3.10, p = .084, \eta^2 = .075$, with no interaction with gender, $F < 1$, suggesting that participants were more accurate on items heard in the familiar voice. There was no interaction between lexical familiarity and talker familiarity, $F < 1$, and no three-way interaction, $F < 1$.

Responses to the unfamiliar talker in the critical conditions were compared with the corresponding items in the control condition. In order to match items in the critical and the control conditions, two separate comparisons were performed—female talker training versus control and male talker training versus control—via Bonferroni-corrected independent sample *t*-tests. There were significantly more harmonic responses for all test conditions. There were significantly more harmonic responses for the male talker training condition, as compared with the control condition,

Table 3 Proportions of same-language, harmonic responses: Means (standard deviations) for all conditions

	Old items		New items	
	Familiar talker	Unfamiliar talker	Familiar talker	Unfamiliar talker
Experiment 1				
Female talker training	.82 (.18)	.70 (.22)	.75 (.20)	.68 (.19)
Male talker training	.88 (.16)	.86 (.20)	.76 (.14)	.74 (.20)
Control	.46 (.12)	.47 (.15)	.52 (.13)	.48 (.14)
Experiment 2				
Female talker training	.84 (.22)	.75 (.11)	.64 (.20)	.73 (.19)
Male talker training	.75 (.19)	.69 (.24)	.63 (.24)	.64 (.20)
Control	.45 (.14)	.51 (.16)	.54 (.11)	.44 (.12)

for new–unfamiliar talker items, $t(50) = 7.94$, $p < 0.0001$, and old–unfamiliar talker items, $t(50) = 5.44$, $p < .0001$, as well as for the comparisons between the female talker condition and the control condition: old–unfamiliar talker items, $t(50) = 6.56$, $p < .0001$; new–unfamiliar talker items, $t(50) = 3.72$, $p < .001$.

Transfer deficits were detected using planned comparisons between familiar and unfamiliar talkers for old and new items. Male and female talker training conditions were combined to increase power, since there was no difference or interaction with gender in the ANOVA. Means and standard errors are presented in Fig. 1. There was a marginally significant difference between familiar and unfamiliar talkers for old items, $t(39) = 1.96$, $p = .057$ (.85 vs. .80), but there was no significant difference for new items, $t(39) = 0.82$, $p = .42$ (.73 vs. .71). These results are consistent with the hypothesis that talker-specific effects are strongest for known words.

These results suggest that learning a productive phonological pattern takes place at a level at which talker identity does not impede judgments made outside of the lexicon. There was a marginal transfer deficit for old items but no effect for new items,⁵ suggesting that any talker-specific representations are delegated to the lexicon.

Discussion

Learners in the present experiment generalized the harmony pattern to novel words spoken by novel talkers. Previous research (Houston & Jusczyk, 2000) demonstrated that infants are more likely to generalize to a novel talker if they first hear either a familiar word or a familiar talker. It is possible that learners generalized to the novel talker for novel items only because they heard familiar items first. To verify this, we partitioned all new/unfamiliar-talker test items that occurred before any other test item. While many

⁵ Analyses were also conducted on the basis of reaction time, with similar results. Because the task was not a speeded judgment task and involved a binary response, any analysis on reaction time must be taken with extreme caution and are thus not included in the main text.

participants did not hear this type of test item first (i.e., those who heard the familiar talker in the first block), all responses to items in this category were correct (i.e., listeners chose the harmonic response on all trials).⁶

Another factor that may have led to increased generalization to unfamiliar talkers was the instructions given to participants at the time of training. The instructions in Experiment 1 noted that participants “need not memorize” the words that they heard. This instruction may have primed participants to form an abstract phonological rule. To test this possibility, we ran a small set of participants ($n = 10$) with the instructions “please memorize the words you hear.” The overall pattern of results was consistent with the results of Experiment 1.

Experiment 1 demonstrates that adult learners are able to extend a novel phonological pattern to unfamiliar talkers. However, all the talkers in Experiment 1 spoke with an American English accent. It is possible that formation of talker-independent representations is contingent on familiarity with the accent in question. For example, speakers may know where to look for talker-specific effects versus talker-independent effects in English speech and are able to extend the pattern to unfamiliar talkers on the basis of this prior knowledge. Experiment 2 explored this possibility through a replication of Experiment 1, using Turkish talkers and English-speaking participants.

Experiment 2

Experiment 2 was identical to Experiment 1, except that the stimuli made use of Turkish-speaking talkers, rather than English-speaking talkers. Replacing English-speaking talkers with Turkish-speaking talkers increases the ecological validity of the present experiment because learning a

⁶ Because there were a relatively small number of test items (10), we ran a small set of participants ($n = 4$) who heard all 10 new/unfamiliar-talker items before any other item. Each of the 4 participants showed results that were consistent with the results of Experiment 1.

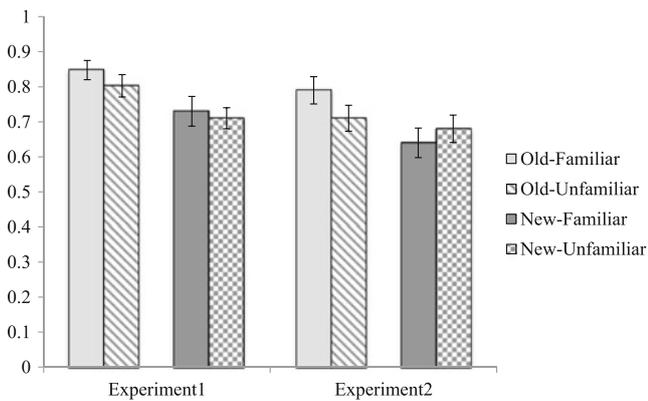


Fig. 1 Transfer deficits for Experiments 1 and 2 (means and standard errors)

novel language requires learning a novel accent. Because Turkish naturally has vowel harmony, speakers hearing harmony in a Turkish dialect are more closely simulating the experience of a second-language learner learning a novel vowel harmony pattern.

Method

Participants

Participants were adult native speakers of English with no knowledge of a vowel harmony language and had not participated in a previous vowel harmony learning experiment. Forty University of Rochester undergraduate students and affiliates were paid \$5–\$10 for their participation. There were 28 participants in the critical condition and 12 participants in the no-training control condition.

Design

The design of Experiment 2 was identical to Experiment 1, except that the talkers used in the experiment were Turkish speakers. Half of the participants were exposed to the female Turkish talker, and the other half were exposed to the male Turkish talker. All participants heard all test items in a random, mixed order. Because the blocked design did not vary, as compared with the random, for simplicity, all participants received the same design.

Materials

Materials were collected in the same manner as in Experiment 1, with a few minor differences. First, two native Turkish speakers from Istanbul recorded the experimental stimuli. All of the vowels and consonants from Experiment 1 are found in Turkish, making it possible to use identical stimuli. However, all languages differ in their pronunciation of vowels

and consonants, meaning that the Turkish stimuli were qualitatively different from the English stimuli. For example, English vowels are often produced as diphthongs, while Turkish vowels are not. While neither of the talkers was a native English speaker, the talkers were fluent in English. Second, the talkers were told to produce the words to represent words spoken in Turkish as closely as possible. In order to create a natural environment for the Turkish speakers, the talkers were asked to “speak as clearly and accurately as possible,” making it difficult to control for length of utterances. Thus, there were differences in durations between the male and the female talkers (with the female talker being slightly longer). Male suffixed items averaged 590 ms, with a range of 463–810 ms.⁷ Female suffixed items averaged 776 ms, with a range of 628–1,033 ms. These differences in length make the talkers even more distinguished. If anything, this difference should prevent generalization to the novel talker.

Procedure

The procedure was identical to that in Experiment 1.

Results

Proportions of same-language, harmonic responses are provided in Table 3. Results are categorized in terms of the control condition and the two critical conditions—male talker training and female talker training—and divided by old items and new items for familiar and unfamiliar talker test items. Control Items have male as the default familiar talker, but statistical comparisons were made according to the appropriate gender.

A mixed-design ANOVA was used to compare responses in the female talker training condition with those in the male talker training condition, using a 2 (gender) × 2 (talker familiarity) × 2 (lexical familiarity) design. There was no effect of gender, $F(1, 26) = 1.26, p = .27, \eta^2 = .046$, suggesting that participants learned the harmony pattern in both experimental conditions. As in Experiment 1, there was a significant effect of lexical familiarity, $F(1, 26) = 6.87, p < .05, \eta^2 = .21$, and no interaction with gender, $F < 1$, suggesting that participants were more accurate with old items than with new items. There was no effect of talker familiarity, $F < 1$, and no interaction with gender, $F < 1$, suggesting that participants were not significantly more accurate on items heard in the familiar voice. This may be due to the fact that participants in both the male talker training condition

⁷ It is unclear whether the differences in speech rate are due to differences between English and Turkish or due to the fact that the talkers in Experiment 1 were more experienced in producing experimental stimuli and thus spoke more carefully.

and female talker training condition were more accurate on old items when they were heard in a familiar voice but were numerically more accurate on new items when they appeared in an unfamiliar voice. This is reflected in the marginal interaction between lexical familiarity and talker familiarity, $F(1, 26) = 3.05$, $p = .092$, $\eta^2 = .10$.

Responses to the unfamiliar talker in the critical conditions were compared with the corresponding items in the control condition. In order to match items in the critical and the control conditions, two separate comparisons were performed—female talker training versus control and male talker training versus control—via Bonferroni-corrected independent sample t -tests. There were significantly more harmonic responses for the male talker training condition than for the control condition for new–unfamiliar items, $t(26) = 4.47$, $p < .05$, and old–unfamiliar items $t(26) = 2.47$, $p < .05$, as well as for the comparisons between the female talker condition and the control condition: old–unfamiliar, $t(22) = 5.64$, $p < .001$; new–unfamiliar, $t(22) = 2.61$, $p < .05$. These results suggest that participants in the critical conditions learned the harmony pattern and extended that pattern to unfamiliar talkers.

Transfer deficits were detected using planned comparisons between familiar and unfamiliar talkers for old and new items. Male and female talker training conditions were combined to increase power, since there was no difference or interaction with gender in the ANOVA. Means and standard errors are presented in Fig. 1. There was a marginally significant difference between familiar and unfamiliar talkers for old items, $t(27) = 1.71$, $p = .099$ (.79 vs. .71), but there was no difference between familiar and unfamiliar talker items for new items, $t(27) = 1.56$, $p = .13$ (.64 vs. .68), and this trend was in the opposite direction of a talker-specific interpretation. These results are consistent in part with the hypothesis that productive linguistic patterns make use of talker-independent representations for novel items but lexical representations are more likely to make use of fine-grained talker-specific representations.

The results of Experiment 2 parallel those of Experiment 1. This suggests that the ability to form a general phonological pattern is not contingent on familiarity with the accent of the talker.

Discussion

Overall, the responses in Experiment 2 were slightly less accurate than the responses in Experiment 1 (e.g., the overall average accuracy in Experiment 1 was 77 %, as compared with 70.5 % in Experiment 2, which was marginally significant) $F(1, 66) = 3.08$, $p = .0084$. It likely that this decline in accuracy is due to the fact that learners in Experiment 2 had to cope with a novel accent in the training phase.

General discussion

Participants in the present study learned a novel back/round vowel harmony pattern and extended that harmony pattern to an unfamiliar talker. This occurred for both Turkish- and English-speaking talkers. Marginal transfer deficits occurred for old items, but no deficits appeared for new items, suggesting that talker-independent mechanisms are at work when listeners make novel judgments regarding categorical phonological patterns.

In the present experiments, the talkers shared the same dialect. It is possible that generalization of a novel pattern across speakers of different dialects may be less robust than generalization across the same dialect. This is consistent with the phonetic relevance hypothesis (Sommers & Barcroft, 2006), which states that the elements relevant to the phonetic code are the ones that listeners will apply to multiple talkers. In natural learning situations, children are exposed to multiple talkers but will typically be exposed to a single primary talker, such as the primary caregiver. Young infants are able to generalize properties of speech from one speaker to another, so long as the talkers' voices are relatively similar (Houston & Jusczyk, 2000; Schmale & Seidl, 2009). Schmale and Seidl showed that 9-month-old infants could accommodate different speakers from their native dialect but could no longer do so when the novel talker spoke in an unfamiliar dialect. One goal for future research is to explore how the features of multiple talkers during exposure affect extension to unfamiliar talkers at test.

The vowel harmony pattern in the present study was relatively simple, including only four vowels. This raises the possibility that learners did not form a productive harmony pattern, but a simple association between stem and suffix vowels. This interpretation is unlikely because prior research has demonstrated that learners of a novel vowel harmony pattern are able to generalize to novel vowels outside a four-vowel inventory (Finley & Badecker, 2009a). Furthermore, learning a novel vowel harmony pattern decreases when the associations between stem and suffix vowels are arbitrary (Pycha et al., 2003). Finally, the categorical nature of the pattern used in the present study holds despite its simplicity.

The present study extends prior research demonstrating that novel phonological patterns can be learned very rapidly, since participants were given only about 10 min of exposure. This raises the possibility that learners did not have enough time to learn the talker's idiosyncrasies, resulting in minimal transfer deficits. While this is an important possibility, the goal of the present study was to assess whether learning categorical phonological patterns requires talker-specific representations. Because the harmony pattern was successfully learned in a matter of minutes, either learning does not require a strong sense of familiarity with the talker, or

the idiosyncrasies of the talker were learned rapidly. Learners may have become familiar with the talker very quickly because only a single talker was heard during training. It is unlikely, however, that the use of a single talker led the learner to assume that all aspects of the talker's speech were general to the language at hand. Previous studies have shown that generalization to novel talkers increases with the number of talkers heard during training (Magnuson et al., 1995).

Categorical, exceptionless patterns may be more susceptible to talker-independent representations than fine-grained phonetic patterns (Lively et al., 1992; Logan et al., 1991; Magnuson et al., 1995). Phonetic patterns tend to be more continuous and variable in terms of rate and consistency of application. Phonological patterns tend to be described in terms of categorical features and segments and exhibit lower levels of variability. While categorical linguistic patterns are subject to exceptions, these exceptions tend to be principled (Zonnefeld, 1978). An important question for future research would be to understand the cases, if any, where talker-specific details play a role in judgments for novel instances of a categorical phonological pattern. One possibility is that learners rely more on talker-specific details when the phonological pattern shows high levels of exceptions or lexical constraints. Another possibility is that using a task that orients the learner toward the phonetic details of the pattern, such as the tasks used in previous studies (Magnuson & Yamada, 1995; Nygaard & Pisoni, 1998), may yield more talker-specific responses. One goal of future research is to determine both the properties of the pattern and the properties of the tasks that create an environment where learners are more prone to respond with respect to talker-specific details.

There is evidence that talker-specific details found in phonetic patterns are stored in the lexicon (Goldinger, 1998) and are available during lexical access (Salverda et al., 2007). Lexical items may be subject to talker-specific details even for words that are learned in a short amount of time and have no semantic content. This predicts that if the vowel harmony pattern in the present study are subject to greater phonetic or lexical variability, generalization to unfamiliar talkers may decline. Learners in the present experiments showed a marginally significant transfer deficit for old items, but not new items. This supports the hypothesis that learning novel categorical patterns involves multiple levels of representations (Luce, Goldinger, Auer, & Vitevitch, 2000; Luce & Large, 2001). Understanding how talker-specific details are used at various levels of representation may shed light on the mechanisms required to integrate productive patterns into an exemplar model of cognition (Goldinger, 1996, 1998; Nosofsky, 1988). Creating a theory that allows for various talker-specific effects at the phonetic level and the lexical level, as well as the categorical level, may lead the way to a better understanding

of the interaction between low-level speech processes, categorical phonological patterns, and the lexicon (Pierrehumbert, 2001, 2003).

While talker-specific effects are clearly language specific, there are important parallels with other areas of cognitive science. The study of learning and generalization across novel items has important consequences for understanding the mechanisms that underlie learning and generalization (Dienes & Altmann, 1997). These consequences apply to both linguistic and nonlinguistic cognitive processes. For example, object recognition and category discrimination require the ability to distinguish between individual exemplars, but also the ability to generalize to novel instances (Nosofsky, 1988). Understanding how different patterns in cognition are subject to different levels of representation and generalization may help to create a unified theory of cognition.

The present study tested the role of talker-specific representations in learning novel phonological patterns. Participants were exposed to a novel vowel harmony pattern in the voice of a single talker and then were tested on both a familiar and an unfamiliar talker. Participants generalized to novel talkers, supporting the hypothesis that learners make use of abstract representations in making judgments regarding novel categorical phonological patterns. Learners showed a marginal transfer deficit for old items, but no deficits for new items, supporting a theory that learning mechanisms for phonological patterns make use of representations that allow for generalization beyond the familiar.

Appendix

Table 4 Full stimuli lists

Training items	Old test items	New test items
budok, budoko	degibe, *degibo	*bipido, bipide
digib, digibe	*giteko, giteke	*mukobe, mukobo
dupob, dupobo	*budoke, budoko	tidipe, *tidipo
gemit, gemite	*dupobe, dupobo	*toguke, toguko
gitek, giteke	gemite, *gemito	*bidito, bidite
kimet, kimete	kukopo, *kukope	butoko, *butoke
kukop, kukopo	monuko, *monuke	godomo, *godome
mekin, mekine	netepe, *netepo	kukogo, *kukoge
midik, midike	*nopube, nopubo	nopuko, *nopuke
monuk, monuko	*tikepo, tikepe	*pedebo, pedebe
netep, netepe		
nopub, nopubo		
puduk, puduko		
tikep, tikepe		
todup, todupo		
tokot, tokoto		

References

- Boersma, P., & Weenink. (2005). Praat: Doing phonetics by computer. Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers*, 25, 257–271.
- Connine, C., & Pinnow, E. (2006). Phonological variation in spoken word recognition: Episodes and abstractions. *The Linguistic Review*, 23, 235–245.
- Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, 106, 633–644.
- Creel, S. C., & Tumlin, M. A. (2009). Talker variability is intrinsic to word representations: Evidence from on-line processing of spoken words. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual Cognitive Science Conference* (pp. 845–850). Austin, TX: Cognitive Science Society.
- Dienes, Z., & Altmann, G. (1997). Transfer of implicit knowledge across domains? How implicit and how abstract? In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 107–123). Oxford: Oxford University Press.
- Finley, S., & Badecker, W. (2008). Analytic biases for vowel harmony languages. *WCCFL*, 27, 168–176.
- Finley, S., & Badecker, W. (2009a). Artificial grammar learning, and feature-based generalization. *Journal of Memory and Language*, 61, 423–437.
- Finley, S., & Badecker, W. (2009b). Right-to-left biases for vowel harmony: Evidence from artificial grammar. In A. Shadl, M. Walkow & M. Abdurrahman (Eds.), *Proceedings of the 38th North East Linguistic Society Annual Meeting* (Vol. 1, pp. 269–282).
- Finley, S., & Badecker, W. (2010). Linguistic and non-linguistic influences on learning biases for vowel harmony. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 706–711). Austin, TX: Cognitive Science Society.
- Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 1166–1183.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105(2), 251–279.
- Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), 1570–1582.
- Johnson, K. (1997). Speech perception without speaker normalization. In K. Johnson & J. W. Mullenix (Eds.), *Talker variability in speech processing* (pp. 145–165). San Diego: Academic Press.
- Koo, H., & Cole, J. (2006). On learnability and naturalness as constraints on phonological grammar. In A. Botinis (Ed.), *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics* (pp. 174–177). Athens.
- Kraljic, T., & Samuel, A. (2005). Perceptual learning in speech: Is there a return to normal? *Cognitive Psychology*, 51, 141–178.
- Kraljic, T., & Samuel, A. (2006). Generalization in perceptual learning of speech. *Psychonomic Bulletin & Review*, 13, 262–268.
- Kraljic, T., & Samuel, A. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, 56, 1–15.
- Kraljic, T., Samuel, A., & Brennan, S. (2008). First impressions and last resorts: How listeners adjust to speaker variability. *Psychological Science*, 19(4), 332–228.
- Lively, S. E., Pisoni, D. B., & Logan, J. S. (1992). Some effects of training Japanese listeners to identify English /r/ and /l/. In Y. I. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.), *Speech perception, production and linguistic structure* (pp. 175–196). Burke, VA: IOS Press.
- Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, 89(2), 874–886.
- Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics*, 62, 615–625.
- Luce, P. A., & Large, N. R. (2001). Phonotaxis, density, and entropy in spoken word recognition. *Language and Cognitive Processes*, 16(5/6), 565–581.
- Magnuson, J. S., & Yamada, R. A. (1995). The effects of talker variability on the acquisition of non-native speech contrasts. *Proceedings of the 1995 International Congress of Phonetic Sciences* (pp. 306–309).
- Magnuson, J. S., Yamada, R. A., Tohkura, Y. i., Bradlow, A. R., Lively, S. E., & Pisoni, D. B. (1995). The role of talker variability in non-native phoneme training. *Proceedings of the 1995 Spring Meeting of the Acoustical Society of Japan* (pp. 393–394).
- Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, 25, 83–127.
- Nosofsky, R. (1988). Exemplar-based accounts of relations between classification, recognition and typicality. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(4), 700–708.
- Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, 60(3), 355–376.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5(1), 42–46.
- Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19(2), 309–328.
- Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. L. Bybee & P. Hopper (Eds.), *Frequency effects and emergent grammar* (pp. 137–157). Amsterdam: John Benjamins.
- Pierrehumbert, J. (2003). Probabilistic phonology: Discrimination and robustness. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 177–228). Cambridge, MA: The MIT Press.
- Pisoni, D. B. (1997). Some thoughts on ‘normalization’ in speech perception. In K. Johnson & J. Mullenix (Eds.), *Talker variability in speech perception* (pp. 9–32). San Diego: Academic Press.
- Port, R. F., & Leary, A. P. (2005). Against formal phonology. *Language*, 81, 927–964.
- Pycha, A., Nowak, P., Shin, E., & Shosted, R. (2003). Phonological rule-learning and its implications for a theory of vowel harmony. *WCCFL*, 22, 101–113.
- Salverda, A. P., Dahan, D., Tanenhaus, M. K., Crosswhite, K., Masharov, M., & McDonough, J. (2007). Effects of prosodically modulated sub-phonetic variation on lexical competition. *Cognition*, 105, 466–476.
- Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, 12(4), 583–601.
- Skoruppa, K., & Peperkamp, S. (2011). Adaptation to novel accents: Feature-based learning in context-sensitive phonological regularities. *Cognitive Science*, 35(2), 348–366.
- Sommers, M., & Barcroft, J. (2006). Stimulus variability and the phonetic relevance hypothesis: Effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification. *Journal of the Acoustical Society of America*, 119(4), 2406–2416.
- Wedel, A. (2006). Exemplar models and language change. *The Linguistic Review*, 24, 147–185.
- Zonnefeld, W. (1978). *A Formal Theory of Exceptions in Generative Phonology*: Walter de Gruyter.