

Constraints on Abstraction: Generalization Across Languages

Sara Finley (finleysr@plu.edu)

Department of Psychology, Pacific Lutheran University
Tacoma, WA, 98447, USA

Abstract

Recent research has suggested a division between lexical representations and phonological patterns; lexical items are stored with talker-specific information, while phonological patterns are represented at a separate, abstract level of representation (Finley, 2013; Smolensky & Legendre, 2006). The present paper provides further evidence for this proposal, demonstrating that learners will extend a novel phonological pattern (vowel harmony) to speakers of a novel dialect when the words are familiar, but not when the words are unfamiliar, further supporting a distinction between the representations of lexical items and the representations of phonological patterns.

Keywords: statistical learning, vowel harmony, phonological representations, lexical representations.

Introduction

One of the major issues in the cognitive science of language is how the phonological representations that make up words are stored in the human mind. Specifically, there is a debate between whether the phonological representations that govern word formation are represented in the mind in terms of abstract, rule-governed processes, or as exemplars of specific instances, based on the phonetic forms of lexical items. While classic phonological theory has proposed highly abstract, symbolic representations for phonological patterns (Chomsky & Halle, 1968), experimental research has questioned the abstract nature of phonological processes, and suggested that phonological rules are epiphenomenal (Port & Leary, 2005). Because listeners encode highly detailed information from the individual speakers they encounter (Nygaard & Pisoni, 1998), researchers have proposed that representations of lexical items are not based on normalized or abstract representations, but based on individual, stored exemplars (Johnson, 1997). Because the generalizations that can be made about phonological restrictions for word formation (e.g., phonotactics and morphophonology) necessarily come from examination of lexical items, it is reasonable to assume that these phonological restrictions share the same constraints on representations as lexical items (i.e., fine-grained, talker-specific representations) (Pierrehumbert, 2001). However, it is also possible that language makes use of multiple levels of representation, with fine-grained exemplar-based representations at one level, and abstract representations at another, higher level (Smolensky & Legendre, 2006).

One of the difficulties in distinguishing between an abstract model of phonological representations and an exemplar model of phonological representations is in teasing apart the representations of individual words and the representations of phonological patterns, since phonological patterns are instantiated through word formation. In order to tease apart the distinction between lexical representations and phonological representations, the researcher must look to how the language handles phonological patterns in novel words. While researchers have investigated questions of phonological representations using a phonological version of a *wug* test (Berko, 1958), these investigations typically assume an abstract, generative view of phonological representations (Becker, Ketrez, & Nevins, 2011; Becker, Nevins, & Levine, 2012; Gouskova & Becker, 2013), and therefore do not address the question of a distinction between lexical representations and phonological rules.

Another approach to testing the differences between lexical items and phonological patterns is by testing how learners treat novel and familiar forms following a brief exposure to a novel language. If there is a distinction between lexical representations and abstract phonological patterns, then learners should treat familiar, known words differently than unfamiliar, unknown words.

Recent research used learning to test the hypothesis that lexical representations and abstract phonological patterns are stored under distinct representations (Finley, 2013). Participants in an artificial grammar learning experiment were exposed to a novel phonological pattern (vowel harmony, where vowels in a word shared the same value of a phonological feature, back and round), and then were tested on their knowledge of that phonological pattern using both familiar words and unfamiliar, novel words. In addition, participants were tested on the same set of familiar and novel items, but these items were spoken by an unfamiliar talker of the opposite gender. Participants were able to extend the newly learned vowel harmony pattern to both novel and familiar items, for both novel and familiar talkers. In addition, the analyses tested for a 'transfer deficit' from familiar to novel talkers. The more speakers make use of talker-specific representations in learning, the larger the transfer deficit should be when accepting items spoken in an unfamiliar voice. There was a clear transfer deficit for familiar items, but not for novel items. This difference suggests that learners store known items in terms of talker-specific phonetic details, but make use of abstract phonological patterns when making grammaticality judgments for novel words. These abstract representations are independent of the specific talker heard during training.

The ability generalize a newly learned phonological pattern to unfamiliar talkers for both familiar and novel

items raises an important question of whether participants in an artificial language learning task generalize to novel talkers more so than they would in a natural setting. In addition, there is a question of whether learners appropriately apply the phonological pattern to the language in question, or whether they will assume that any novel language would follow the constraints of the newly learned phonological pattern, without any positive evidence. The present study addresses this concern, and provides further evidence for distinction between representations for known/familiar words and abstract phonological patterns.

If learners represent the phonological components of lexical items and phonological patterns separately, one should expect this difference to manifest itself when learners are asked to extend a novel pattern to a novel, unfamiliar dialect. It is common for languages to share or borrow lexical items (e.g., ‘computer’) without sharing phonotactic restrictions (Ito & Mester, 1995). This means that the phonological component of a lexical item can be shared from language to language (so long as the phonological restrictions of the borrowing language allow for the phonological form found in the borrowed lexical item), without affecting the grammar of the borrowing language. However, in order to borrow a phonological pattern that affects novel words, the grammar of the borrowing language must be changed. If learners freely borrow the phonological form of known words, but not the phonological pattern that governs the formation, then we should expect generalization of known (familiar) words to an unfamiliar talker of a novel dialect, but not generalization of novel (unfamiliar) words. This predicts that learners of a novel phonological pattern will be able to extend a newly learned phonological pattern to familiar items spoken in an unfamiliar dialect, but not novel items spoken in an unfamiliar dialect. This is essentially the opposite pattern of results that Finley (2013) found. Because Finley (2013) used the same dialect across talkers, speakers were able to freely generalize to the novel talker.

By demonstrating that learners extend a phonological pattern unfamiliar talkers differently depending on the spoken dialect of the talker, it will shed light on the nature of representations for phonological patterns. Specifically, if learners extend the phonological pattern to both familiar and unfamiliar words to novel talkers of an unfamiliar dialect, it suggests that learners are biased to generalize phonological patterns across languages. If learners extend the phonological pattern to novel talkers of an unfamiliar dialect for unfamiliar words but not familiar/known words, it suggests that learners view the representation of the familiar words as talker-specific, but are biased to generalize the novel phonological pattern across languages. If learners extend the phonological pattern to novel talkers of an unfamiliar dialect for familiar words, but not unfamiliar words, it suggests that learners are biased to extend the phonological representation of known words to novel languages, but are biased against applying phonological patterns to a novel language.

Methods

In the present study, learners were trained on a novel phonological pattern, specifically vowel harmony. Vowel harmony is a phonological pattern in which adjacent vowels (ignoring consonants) must share the same value of a phonological feature (Clements, 1976). For example, in a left-to-right back/round vowel harmony system, words that begin with a back/round vowel (e.g., [o], [u]) must only contain back/round vowels, and words that begin with a vowel that is front/unround (e.g., [i], [e]) must only contain front/unround vowels. Previous research has shown that adult, English speaking participants can learn a back/round harmony pattern with relatively minimal exposure to the pattern, particularly if the pattern is presented as a pseudomorphophonological alternation (e.g., as pairs of words, /bodo/-/bodomu/, /bede/-/bedemi, etc.) (Finley, 2013; Finley & Badecker, 2009).

Participants were trained on a vowel harmony pattern spoken by a single talker, and then tested on items spoken by a familiar talker as well as an unfamiliar talker. The unfamiliar talker spoke a distinctly different dialect from the familiar talker (generalizing from Turkish to English or vice versa).

Participants

All 36 participants were adult native English speakers recruited from Waldorf College, a small liberal arts college in Northern Iowa, USA. Each participant was given extra credit in a psychology course for participating. Some participants may have previously participated in an artificial grammar learning experiment. No participant had any previous experience with a vowel harmony system, natural or artificial.

Design

Adult, English speaking participants were exposed to a back/round vowel harmony pattern that was presented in one of two distinct voices: a male native Turkish speaker with an Istanbul accent, or a male native English speaker with an American accent. All exposure items were identical (except for the talker). The harmony pattern was presented to learners in pairs of words that contained a ‘stem’ followed by its ‘suffixed’ form (though participants were only told that they would hear words in pairs). Stems triggered a suffix vowel that was either /-e/ or /-o/ depending on whether the vowels in the stems contained front vowels (/i/ or /e/) or back vowels (/o/ or /u/). All stems were of the form CVCVC (e.g., /betig/) with the vowels following back/round harmony constraints (all stem vowels were either both front or both back, and never disharmonic), and the consonants drawn from the set (/p, t, k, b, d, g, m, n/). There were 24 items, presented as pairs: stem followed by stem+suffix (e.g., /betig, betige/). These 24 items were presented eight times, each in a random order (with suffixed items always following bare stem items). Examples of the exposure stimuli can be found in Table 1.

Table 1: Examples of Exposure Stimuli.

Stem Vowel	Stem	Stem+Suffix
Back	budok	budoko
	nopub	nopubo
	dupob	dupobo
Front	tikep	tikepe
	gemit	gemite
	degib	degibe

Following exposure, participants were presented with a two-alternative, forced-choice test in which participants were asked to decide which was more likely to come from the language they had just heard. Participants heard two CVCVC-V words, each identical except for the final vowel ending: either /-e/ or /-o/. Because the first two vowels obeyed harmony, the choice of the final vowel (/e/ or /o/) depended on the back/round features of the vowels in the first two syllables. The test items were identical for all participants, and included items spoken by the Turkish talker and items spoken by the American talker, presented in a random order. For each participant, half of the items were presented in the voice of a familiar talker, and the other half of the items were presented in the voice of an unfamiliar talker. In addition to items that were heard during the exposure phase, novel items were also presented to participants. This created four types of test items, 10 of each type, creating 40 total test items. Old Items-Familiar Talker were items that were identical to the training set items, and therefore were spoken by a familiar talker. New Items-Familiar Talker were items that were not in the training set, but were spoken by the same talker that was heard during training. Old Items-New Dialect, were items that were heard in the training set, but the talker was unfamiliar, and spoke a dialect that was not heard during the exposure session. New Items-New Dialect items were not heard in the training set, and the talker was unfamiliar, and spoke a dialect that was not heard during the exposure session. Examples of these test items can be found in Table 2.

All stimuli were recorded in a sound attenuated booth at 12,000 Hz. Stress was placed on the first syllable using English or Turkish pronunciation, with the exception that no vowels were reduced, meaning that some English syllables contained partial stress (as English reduces unstressed syllables). All stimuli items were normalized for intensity (set at 70dB), though participants were allowed to adjust headphones to a comfortable volume during the experiment. All stimuli creation and modifications to sound files was performed in Praat (Boersma & Weenink, 2005). All phases of the experiment were run in PsyScope X (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were given both written and verbal instructions. The entire experiment took approximately 20 minutes.

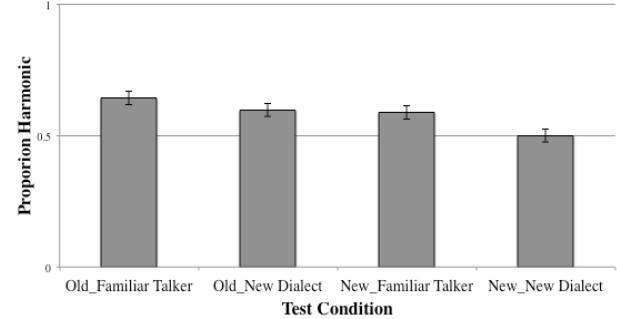
Table 2: Examples of Test Stimuli

Items	Stem Vowel	Harmonic	Disharmonic
Old	Back	budoko	budoke
		nopubo	nopube
		dupobo	dupobe
Front	tikepe	tikepo	
	gemite	gemito	
	degibe	degibo	
New	Back	butoko	butoke
		godomo	godome
		tokugo	tokuge
Front	bedite	bedito	
	bipide	bipido	
	tidipe	tidipo	

Results

Proportion of correct responses for all test items are given in Figure 1. We compared each test item to 50% chance via three separate one-sample *t*-tests.

Figure 1: Results (Means and Standard Errors).



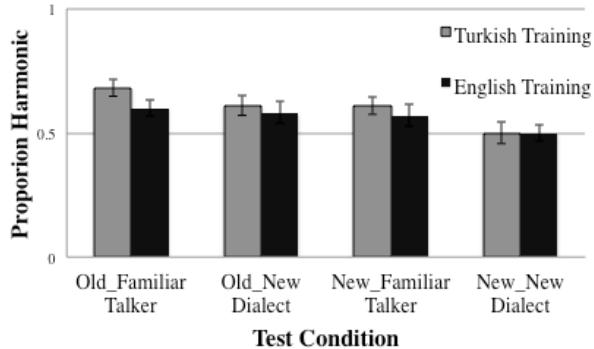
If participants represent individual known items separately from the harmony pattern found in the exposure language, learners will extend the harmony pattern to an unfamiliar talker speaking a novel dialect when the items are familiar, but not when the items are novel. Results support this hypothesis, as represented in Figure 1. Participants were able to extend the harmony pattern to Old Items-Familiar Talker (mean = 0.64, SD = 0.15), $t(35) = 5.67$, $p < 0.001$, New Items-Familiar Talker (mean = 0.59, SD = 0.17), $t(35) = 3.05$, $p < 0.01$, and Old Items-New Dialect items (mean = 0.60, SD = 0.17), $t(35) = 3.34$, $p < 0.01$, but not to Old Items-New Dialect items (mean = 0.50, SD = 0.16), $t(35) = 0.00$, $p = 1.00$. This suggests that participants learned the vowel harmony pattern, but only extended the vowel harmony pattern to a novel dialect when the learner was familiar with the lexical item.

In addition, we performed a 2x2 (Talker x Test Item) generalized linear mixed model fit by the Laplace approximation using the lme4 package in R. A single model was created with random intercepts for both items and subjects. There was a significant difference between both the Old and the New items, $\beta = 0.41$, $z = 2.15$, $p = 0.032$, and the Familiar and New Talker items, $\beta = 0.36$, $z = 2.33$, $p = 0.020$. The interaction between the Talker and Test Item was not significant, $\beta = -0.12$, $z = 0.56$, $p = 0.58$. These results suggest that participants performed better on Old Items and Familiar Talker items, but the transfer deficit between Familiar and New Dialect talkers was not significantly different between Old and New items.

Overall, the results of the present study are distinct from Finley (2013), who showed a transfer deficit for Old items but not for New Items. In fact, the trend is in the reverse direction reverse; there was no significant transfer deficit for Old Items, $t(35) = 1.33$, $p = 0.19$, but there was a significant transfer deficit for New Items, $t(35) = 3.37$, $p < 0.01$.

It is possible that the results were mediated by the familiarity of the dialect. The participants were all American English speakers, and were not familiar with Turkish. Learners may have responded differently to Turkish talkers simply because they were unfamiliar with Turkish. In order to ensure that the results were not due to familiarity with English and unfamiliarity with Turkish, we separated out the Turkish and the English responses, as shown in Figure 2.

Figure 2: Results: Turkish and American Talker Training Separated (Means and Standard Errors)



There was no significant difference between participants who were trained on Turkish compared to participants who were trained on English for Old Items with a Familiar Talker, $t(17) = 1.71$, $p = 0.10$, Old Items with a New Dialect, $t(17) = 0.47$, $p = 0.64$, New Items with a Familiar Talker, $t(17) = 0.49$, $p = 0.63$, or New Items with a New Dialect, $t(17) = 0.00$, $p = 1.0$. This suggests that the response patterns were the same regardless of whether the learner was trained using an English speaking talker or a Turkish speaking talker.

Discussion

Finley (2013) demonstrated that learners can extend a novel phonological pattern (vowel harmony) to a novel talker, for both novel and familiar items. However, the present study

showed a different pattern of results. Participants extended the novel vowel harmony pattern to a novel talker, only when the item was familiar (Old) items. The major difference between Finley (2013) and the present study is that Finley (2013) tested for generalization to a speaker of a different gender (male to female) of the same dialect (Turkish or English). In the present study, the gender was held constant (male in all cases), but the identity and the dialect of the talker changed from Turkish to English and vice versa. The change in results from Finley (2013) to the present study suggests that the representations for phonological patterns are different when extending across genders than when extending across dialects or languages.

One puzzle of the present results is that the overall rate of correct, harmonic responses were relatively low compared to previous artificial grammar learning experiments with a similar paradigm (including Finley, 2013). One possibility for this difference is that the inclusion of the novel dialect made the test more difficult. Because Turkish and English show relatively different vowel acoustics, switching between Turkish and English vowels may have been difficult for participants. The test items varied randomly for talker familiarity, meaning that on one trial, the participant could respond to a Turkish talker, and the next, an English speaking talker. Switching from language to language may have made the task more difficult overall, leading to a decrease in overall harmonic responses. Because the mean for New-New Dialect responses was exactly 0.50 (chance) for both the English talker training and the Turkish talker training, it suggests that the learners were reluctant to extend the vowel harmony pattern to novel talkers when the words were unfamiliar. Of the 36 participants, only nine (25%) showed scores above 50% in the New-New Dialect test items.

Because it is possible that the increased difficulty of the task skewed the results, we re-ran the statistics, excluding all participants who scored at or below chance for New Items from a familiar dialect ($n=14$, 39% of participants), as these participants are least likely to have learned the general vowel harmony pattern. These participants were able to extend the harmony pattern to Old Items-Familiar Talker (mean = 0.67, $SD = 0.17$), $t(21) = 4.87$, $p < 0.001$, New Items-Familiar Talker (mean = 0.69, $SD = 0.19$), $t(21) = 7.27$, $p < 0.001$, and Old Items-New Dialect items (mean = 0.60, $SD = 0.12$), $t(21) = 2.37$, $p < 0.05$, but not to Old Items-New Dialect items (mean = 0.56, $SD = 0.18$), $t(21) = 1.57$, $p = 0.13$. These are the same results as when all data were included.

In addition, we performed a 2x2 (Talker x Test Item) generalized linear mixed model fit by the Laplace approximation using the lme4 package in R. A single model was created with random intercepts for both items and subjects. There was no significant difference between the Old and the New items, $\beta = 0.15$, $z = 0.67$, $p = 0.50$, but there was a significant difference between the Familiar and New Talker items, $\beta = 0.56$, $z = 2.78$, $p = 0.0055$. There was no significant interaction between Talker and Test Item,

$\beta = -0.18$, $z = -0.62$ $p = 0.53$. These results suggest that the effect of Old vs. New test items was carried by the participants who failed to generalize the harmony pattern to New items, but that even participants who successfully extended the harmony pattern to new items still showed a transfer deficit for novel talkers.

A Theory of Phonological Representations

The present data, along with the data from Finley (2013) support a theory of phonological representations in which familiar words (or lexical items) are stored differently from the abstract phonological patterns that govern lexical items. The extension of a phonological pattern depends on whether the pattern is extended to a novel talker of the same dialect, or is a cross-linguistic extension, or borrowing.

When a phonological pattern is learned, the representation of that pattern is stored in terms of the words that instantiate the pattern, as well as the abstract phonological rule, or constraint ranking (Prince & Smolensky, 1993/2004). The abstract representation of the phonological pattern is talker-independent, but language specific.

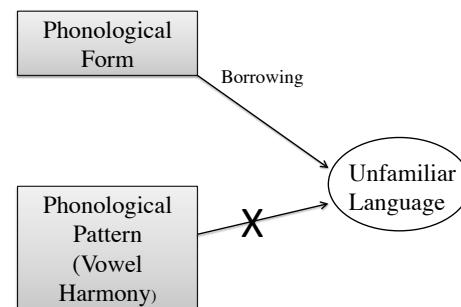
Transfer Across Speakers of the Same Dialect Because the learner stores the phonological pattern in a language-specific manner, the abstract representation of the phonological pattern should apply to all speakers of the language in question. This means that a learner will assume that an unknown speaker of the known language will make use of the phonological pattern in question. For this reason, the phonological pattern is extended to all types of words, Old and New, for both familiar and unfamiliar talkers, so long as the talkers are speaking the same basic dialect. However, because the familiar items are stored in terms of the voice heard during exposure, there should be a transfer deficit for familiar items. Because the novel words are stored abstractly, with a talker independent representation, there will be no transfer deficit. These, essentially, are the results of Finley (2013).

Transfer Across Speakers of a Novel Dialect When the learner is asked to make generalizations about a novel language from speakers of different dialect, a different set of assumptions must be made with respect to when the phonological pattern will apply. If the abstract phonological pattern is stored in a language-specific manner, then the learner will only extend the phonological pattern to speakers of the same dialect that was heard during exposure. For this reason, the learner will not transfer the phonological pattern to speakers of an unfamiliar dialect. However, the lack of transfer should only apply to novel items, because novel (New) items reflect the abstract, language-specific representation of the phonological process, whereas the familiar (Old) items reflect the beginnings of the formation of a lexical item. Because lexical items can be borrowed across languages and dialects, learners can be expected to extend the phonological form of the familiar words to the novel language. In this sense, the learner is extending the

phonological form of known words as if they were lexical items, but is not extending the phonological pattern across speakers of different dialects. This is essentially the result of the present study.

This general proposal is visualized in Figure 3. Old items are represented in terms of their phonological form. This representation is subject to borrowing into a novel, or unfamiliar language. The phonological representation (vowel harmony) that is represented as the New items in the experiment, cannot be borrowed as freely into the novel, unfamiliar dialect.

Figure 3: Separate Representations for Phonological Forms and Phonological Patterns Leads to Familiar Words but not Unfamiliar Words Borrowed to a Novel Language



Learners store both the phonological form of each individual word, as well as the phonological pattern that governs word formation. The phonological form of novel words is governed by abstract rules or constraints that are specific to the language in question.

The above model supports a view of phonological processes in which abstract patterns explain learners' variable generalizations to novel items. While exemplar models of phonological representations have the capacity to from generalizations over a training set (Kirchner, Moore, & Chen, 2010), the generalization must be variable for different types of items— novel items, novel talkers and novel dialects, in order to capture the results found in the present and previous studies. It is unclear whether this type of exemplar model would be more or less complex than a model that assumes abstract representations, or if such an exemplar model would create representations that are indistinguishable from the representations that form abstract phonological patterns.

The artificial grammar learning paradigm used in the present study does not provide semantic representations. Further, with only eight repetitions of the exposure items, it is unlikely that learners formed a fully functional lexicon of the language presented during exposure. Despite this, learners formed a familiarity with the words presented in the training set, enough to show differential treatment of familiar and novel items. This suggests that participants

relied on different strategies for responding at test depending on whether the test item was a familiar word or an unfamiliar word. When the word was familiar, learners could rely on the familiarity of the word (even if only partially), as well as the phonological pattern (vowel harmony) that governed the phonological representation. Unfamiliar, novel words, can only be judged based on the phonological pattern that produces them.

The present study supports a view of phonological representations in which lexical items are represented distinctly from the phonological patterns that govern word formation. In this view, abstract phonological processes are not epiphenomenal, but a separate level of representation from lexical items. This separation can help explain why lexical items are borrowed without close contact, while phonological patterns (and other grammatical patterns) require more intense contact to be borrowed. For a lexical item to be borrowed, the speaker need not change the grammar of the borrowing language, but for a phonological (or other grammatical) process to be borrowed, the speaker must apply the rules that govern word formation from the host language to the borrowing language, a process that would typically require greater contact and shared representations between the two languages.

Conclusions

This study presented an artificial grammar learning experiment in which learners were trained on a novel vowel harmony pattern. Participants were tested on items spoken a the familiar voice, and a novel voice, but with a different dialect. Participants were only able to extend the vowel harmony pattern to the novel dialect when the items were familiar (present in the training set), suggesting distinct representations for abstract phonological processes and stored lexical items.

Acknowledgments

I would like to thank Nathan Fursa for his help creating the stimuli. I am grateful to Ariel Goldberg and Ilker Yildirim for lending me their voices. I am also indebted to Patricia Reeder, Neil Bardhan, and additional members of the ANLab at the University of Rochester for help in designing and implementing this study. I would also like to thank the four anonymous reviewers for helpful comments and insights. I assume responsibility for any and all errors.

References

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

Becker, M., Nevins, A., & Levine, J. (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language*, 88(2), 231-268.

Berko, J. (1958). The child's learning of English morphology. *Word*, 15, 150-177.

Boersma, P., & Weenink. (2005). Praat: Doing phonetics by computer.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper & Row.

Clements, G. N. (1976). The autosegmental treatment of vowel harmony. *Phonologica*, 111-119.

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.

Finley, S. (2013). Generalization to novel talkers in artificial grammar learning. *Psychonomic Bulletin & Review*, 20(4), 780-789.

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

Gouskova, M., & Becker, M. (2013). Nonce words show that Russian yer alternations are governed by the grammar. *Natural Language and Linguistic Theory*, 31(3), 735-765.

Ito, J., & Mester, A. (1995). The core-periphery structure of the lexicon and constraints on reranking. In J. N. Beckman, L. W. Dickey & S. Urbanczyk (Eds.), *Papers in Optimality Theory* (Vol. 18, pp. 181-209). Amherst: GLSA.

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.

Kirchner, R., Moore, R. K., & Chen, T. Y. (2010). Computing phonological generalization over real speech exemplars. *Journal of Phonetics*, 38(4), 540-547.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, 60(3), 355-376.

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.

Port, R. F., & Leary, A. P. (2005). Against formal phonology. *Language*, 81, 927-964.

Prince, A., & Smolensky, P. (1993/2004). *Optimality theory: Constraint interaction in generative grammar*. Cambridge: Blackwell.

Smolensky, P., & Legendre, G. (Eds.). (2006). *The harmonic mind, Volume 1: Cognitive architecture*. Cambridge, MA: MIT Press.