

LEARNING NON-NATIVE PHONOTACTIC CONSTRAINTS OVER THE WEB

Ariana Steele¹, Thomas Denby¹, Chun Chan¹, Matthew Goldrick¹

¹Department of Linguistics, Northwestern University, U.S.A.

arianasteele2016@u.northwestern.edu

ABSTRACT

Although it is well known that a listener's linguistic experience influences their perception of non-native speech sounds, many researchers have access to a limited range of listener populations. Building on recent work, we examined if this issue could be addressed by conducting non-native speech perception experiments over the internet. We developed a web-based version of an implicit learning paradigm that exposes participants to novel phonotactic constraints. We found that native English listeners could acquire novel constraints based on both native- and non-native language auditory stimuli. Critically, the degree of success in learning non-native patterns reflected the relationship between the native- and non-native sound systems. This suggests that web-based experimentation is a viable means of examining non-native speech perception, creating opportunities to test a wider array of participant populations.

Keywords: L2 speech perception; phonotactics; implicit learning; web-based experiments

1. INTRODUCTION

An extensive body of research has shown how perception of non-native speech is influenced by the experience of the listener [1, 2, 3, 4]. A challenge to such research is that many researchers have reliable access to a limited number of listener populations. For example, at many universities in the United States, the vast majority of young adult participants are native English speakers. One possible means of addressing this issue is to take speech perception experiments out of the lab. Recent work suggests that it is possible to gather reliable data from internet-based experiments using auditory stimuli [5]; furthermore, online crowdsourcing marketplaces provide some degree of access to listeners from different language backgrounds [6].

In order to consider the feasibility of this approach, an on-line paradigm was used to assess the perception of native English speakers (who have been extensively studied in the laboratory; e.g., [4]). We examined their perception of non-native French speech, focusing on the acquisition of phonotactic

constraints (i.e., restrictions on sound sequences). We adapted an implicit learning paradigm [7] for use over the web. We first verified that this online paradigm allowed participants to successfully acquire constraints from native language stimuli. We then examined whether participants could learn constraints from non-native speech. Consistent with previous work, participants' acquisition of non-native phonotactic constraints was sensitive to the structure of their native language sound inventory. This suggests that web-based experiments can provide the sensitivity needed to investigate the influence of listener background on non-native speech perception. We conclude by discussing the promises and challenges of such experiments.

2. PHONOTACTICS AND SOUND CATEGORY CONTRAST IN L2 PERCEPTION

Perception of non-native speech sounds reflects, in part, the relationship between non-native and native sound contrasts ([1, 2, 3]). For example, in French, syllables like /si/, /su/, and /sy/ form minimal pairs. However, in English, only /si/ and /su/ form a minimal pair; /sy/ is perceived as a variant of /su/. This influences how English speakers perceive French sounds. Because English has only one high rounded vowel category, /u/, English speakers often have difficulty distinguishing French /y/ from /u/. In contrast, the distinction between French /y/ and /i/ is less difficult; English speakers perceive these as two different categories ([4]).

The perception of non-native speech sounds can also be affected by phonotactic constraints of the native language. For example, a phonotactic constraint of Japanese requires the presence of a vowel between two word-medial obstruents. Japanese listeners thus have difficulty distinguishing between stimuli where such vowels are present vs. absent [8]. The current study examines a complementary question: how does the ability to perceive non-native sounds influence the acquisition of phonotactic constraints?

In order to learn *how* phonemes are distributed in a language (i.e. phonotactics), speakers must first know *which* sounds are contrastive in the language. Phonotactic constraints contingent on easy-to-perceive non-native contrasts (e.g., /i/-/y/ for English

listeners) should therefore be more robustly acquired than constraints dependent on difficult-to-perceive contrasts (e.g., /y/-/u/ for English listeners).

3. IMPLICIT PHONOTACTIC LEARNING

This clear prediction provides an ideal testing ground to examine the efficacy of web-based studies of non-native perception. Bernard and colleagues ([7, 9]) developed a lab-based recognition memory paradigm in which participants implicitly learn phonotactic constraints. Participants first hear multiple repetitions of *familiarization* syllables that obey an experiment-specific phonotactic constraint. After the presentation of each syllable, participants indicate whether they have previously heard that syllable during the experiment. After this initial phase, they continue to repeatedly hear the familiarization syllables randomly intermixed with novel *test* syllables that either respect or violate the constraint. Due to implicit learning of the phonotactic constraint, participants make consistent errors on the test syllables. The legal test syllables fit the pattern established by the familiarization syllables, but the illegal ones do not. Participants are thus more likely to mistakenly recall legal vs. illegal novel test syllables as having been presented on previous trials [7].

Our first goal was to assess whether this paradigm could be adapted as a web-based experiment. In Experiment 1 (‘native, two-category’), native English listeners heard monosyllables consisting of native-language segments. In the familiarization syllables, the manner of the coda consonant (stop vs. fricative) was determined by the preceding vowel (/i/ vs. /u/, counterbalanced across participants). Previous work has shown that similar constraints can reliably be acquired by adult [9] and infant [10] participants.

We then extended the paradigm to non-native speech. Experiment 2 (‘non-native, two-category’) investigated a phonotactic constraint contingent on a non-native vowel distinction that maps onto two categories in English (French /i/-/y/, mapping onto English /i/-/u/). Experiment 3 (‘non-native, one-category’) examined learning under conditions predicted to be more difficult: a non-native distinction mapping onto a single English category (French /y/-/u/, mapping onto English /u/).

4. EXPERIMENTAL METHODS

4.1. Participants

Ninety one participants were recruited through Amazon Mechanical Turk (AMT; receiving \$3 as

compensation for their time, the same hourly rate as in-lab participants). In order to incorporate only those participants who clearly attended to the task, we only included data from participants who correctly answered ‘yes’ to familiarization stimuli at a rate of at least 90% and correctly answered ‘no’ to test stimuli at least 10% of the time. We tested 16 such participants in each experiment, for a total of 48 across the 3 experiments. While the exclusion rate was high, conducting each experiment took only a few hours’ time; thus, the drop-out rate did not have significant consequences in terms of experimenter resources.

All participants were self-identified native English speakers. Six participants reported some knowledge of French. The results were qualitatively similar when these participants were excluded.

4.2. Acoustic Stimuli

For Experiment 1, a set of 36 syllables were recorded at a 44.1 kHz sampling rate by a native English speaker. Stimuli were consonant-vowel-consonant syllables with voiceless stops /p, t, k/ and voiceless fricatives /f, s, ʃ/ as onsets. Vowels were either /i/ or /u/. Codas were either a voiceless stop or a voiceless fricative drawn from the same sets as the onsets. Experiments 2 and 3 stimuli were recorded by a native French speaker naïve to the experimental design. The structure of these syllables was identical to Experiment 1 except English /u/ and /i/ were replaced with French /y/ and /i/ for Experiment 2 and /y/ and /u/ for Experiment 3. All stimuli were normalized to 0.06 Pa.

4.3. Design

Each experiment relied on the vowel restricting the manner of the coda (fricative or stop). The syllables were divided into two complementary groups: familiarization and test. Familiarization syllables respected a particular phonotactic constraint in order to familiarize participants during the first stage of the experiment. In the second stage of the experiment, test syllables were presented in addition to familiarization syllables. Half of the test syllables respected the constraint seen in the familiarization syllables and half violated it.

The 18 familiarization syllables displayed one of two vowel-contingent phonotactic constraints: in Experiment 1, for example, either /u/→stop and /i/→fricative *or* /i/→stop and /u/→fricative. Test syllables followed the phonotactic rule but were not presented as familiarization stimuli (test-legal

syllables) as well as novel syllables that violated the phonotactic rule (test-illegal syllables).

Each participant was randomly assigned to one of two groups based on phonotactic constraint and counterbalanced to evenly test all rules: Group A (/i/→fricative and /u/→stop) or Group B (/u/→fricative and /i/→stop). The assignment of items to familiarization vs. test sets was counterbalanced within each group.

4.4. Procedure

Participants were recruited through AMT and referred to the experimental website hosted on a Northwestern University server. After completion of the experiment, participants received compensation through AMT.

Participants first electronically signed a consent form, after which they were given an audio test. Participants listened to an English word and had to type the word correctly to continue. This process was repeated for a second word. This ensured that participants' speakers functioned correctly and so that participants could adjust their volume to comfortable levels. Participants could play the word as many times as needed. This was followed by a brief demographic questionnaire based on those regularly administered in the laboratory setting, concerning places of long residence, second and third language experience, dialect information, and speech or hearing impairments.

Experimental trials were divided into six blocks; however, participants were neither told nor shown a distinction between blocks. Participants heard the 18 familiarization syllables repeated in random order twice in all six blocks. During familiarization (Blocks 1 and 2), participants were presented with only syllables that respected the experiment-specific phonotactic constraint. During generalization (Blocks 3-6), in addition to hearing the same 18 familiarization syllables presented twice each, participants were presented with 36 test syllables. The test syllables were each presented once, with 4-5 of both test-legal and test-illegal syllables per block, counterbalanced and randomly intermixed with the familiarization syllables. This resulted in 252 stimuli over the six blocks: 18 familiarization syllables presented twice per block, as well as 18 test-legal and 18 test-illegal syllables presented once each.

After the presentation of each syllable, participants answered the question, "Have you heard this sound previously in the experiment?" by clicking either the "Yes" or "No" button on the screen. No feedback was provided.

5. RESULTS

This paradigm provides two measures of learning. First is retention of the familiarization stimuli, shown in Table 1. Here and below, 95% confidence intervals (CIs) for means were estimated via bootstrap resampling with 1,000 replicates (in this analysis, the distribution of a statistic is estimated by re-sampling from the observations with replacement). This shows that participants in each condition had comparable retention of the training data; therefore, any differences on test items reflect differences in participants' ability to generalize what they have learned to novel syllables.

Table 1: Mean percentage of 'yes' responses to familiarization syllables (95% CIs shown in parentheses).

Experiment	Percentage 'Yes' on Familiarization Syllables
Native, two-category	97.0% [95.4%, 98.4%]
Non-native, two-category	98.8% [98.3%, 99.4%]
Non-native, one-category	97.4% [96.2%, 98.6%]

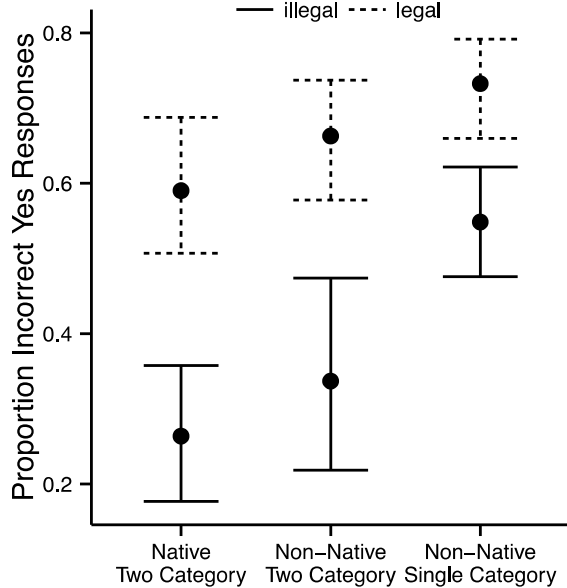
The second measure of learning is generalization of the phonotactic constraint to the novel test syllables (reflected by mistaken recall of syllables that obey the constraint). The likelihood that participants would respond 'yes' incorrectly was analysed using a linear mixed effects regression model. This included a contrast-coded predictor of test syllable type (illegal vs. legal) with uncorrelated random intercepts and slopes by participants and items. (Models with correlated random slopes failed to converge in some analyses and were therefore omitted throughout.) Significance was assessed via model comparison [11].

As shown in Figure 1, in the native, two-category experiment, participants successfully generalized the novel phonotactic constraint. They were more likely to incorrectly respond 'yes' to legal vs. illegal test syllables ($\beta = 3.9$, s.e. $\beta = 0.20$, $\chi^2(1) = 54.15$, $p < 0.0001$), reflecting their implicit learning of the pattern present in the familiarization syllables. This shows that the implicit learning paradigm can be extended to web-based presentation.

Experiments 2 and 3 revealed that participants can also generalize novel phonotactic constraints based on non-native distinctions. In the non-native two-category experiment (Experiment 2),

participants were again more likely to incorrectly respond ‘yes’ to legal vs. illegal test syllables ($\beta = 3.9$, s.e. $\beta = 0.20$, $\chi^2(1) = 45.05$, $p < 0.0001$). Participants also successfully acquired the constraint in the non-native, single category experiment ($\beta = 2.8$, s.e. $\beta = 0.20$, $\chi^2(1) = 43.44$, $p < 0.0001$).

Figure 1: Mean proportion of incorrect “yes” responses for the three experiments for illegal (solid) vs. legal (dashed) test syllables. Error bars show 95% CIs.



Although participants successfully generalized in all three experiments, the degree to which participants differentiated legal vs. illegal test syllables was clearly smaller in the non-native single category condition. This was confirmed by follow-up regressions. These extended the regressions above by including an additional contrast-coded factor reflecting the different experiments. This factor interacted with legality; a significant interaction would indicate that participants in one condition did not distinguish legal vs. illegal syllables to the same degree as in the other condition.

This interaction was significant in regressions comparing the non-native single category experiment to both the native ($\beta = -1.01$, s.e. $\beta = 0.27$, $\chi^2(1) = 11.92$, $p < 0.001$) and non-native two-category experiments ($\beta = -1.04$, s.e. $\beta = 0.28$, $\chi^2(1) = 11.01$, $p < 0.001$). Participants in the single-category experiment showed less sensitivity to the legal vs. illegal test syllable distinction than participants in two-category experiment. There was no reliable difference between the non-native vs. native two-category experiment ($\chi^2(1) < 0.1$, $p > 0.85$).

6. DISCUSSION

Research into non-native speech perception has been hampered by the difficulty of accessing a range of listener groups. Our results suggest that web-based experimentation can plausibly provide one means of addressing this issue. We adapted an implicit learning paradigm used to study phonotactic learning for web-based presentation. Participants successfully generalized phonotactic constraints based on native and non-native stimuli. Furthermore, performance was sensitive to the language experience of the listeners. Participants exhibited weaker generalization of phonotactic constraints based on a difficult- vs. easy-to-perceive non-native vowel contrast.

Additional work is required to determine if this approach provides a comprehensive solution to gathering a diverse sample of listener backgrounds, such as native speakers of languages other than English. To what degree do internet-based crowdsourcing marketplaces provide access to a good sample of different listener groups? While AMT provides some degree of access [6], its diversity is limited. This suggests that researchers may need to develop other means of recruiting diverse populations over the web.

Another general issue is that many lab-based paradigms might not be optimally presented over the web. Web-based presentation allows for a wide array of possible distractions and variation in the experience of participants. This likely contributed to the high rate of exclusion (47%) of participants in our experiment. We addressed this issue by simply running more participants, excluding those that failed to meet our performance criteria. However, this biased sampling of participants might skew some results. An alternative approach would be to develop more engaging experimental paradigms; for example, making experiments more game-like (e.g., [12]).

In spite of these challenges, we were able to efficiently and effectively investigate the effect of non-native sound categories on implicit phonotactic learning outside of the lab. As predicted, the degree of learning corresponded to differences between native and non-native sound category structure. This creates new opportunities for examining speech perception in populations beyond the samples of convenience used in laboratory studies.

7. REFERENCES

- [1] Best, C. T. (1995). A direct-realist view of cross-language speech perception. In W. Strange (Ed.),

Speech perception and linguistic experience: Issues in cross-language speech research (pp.).
Timonium, MD: York, 171-206.

- [2] Flege, J. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (Ed.) *Speech perception and linguistic experience: theoretical and methodological issues* Timonium, MD: York, 229-273.
- [3] Kuhl, P. K. (2000). Phonetic learning: a pathway to language. *Proc. Nat. Acad. Sci.* 97, 11850-11857.
- [4] Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *J. Acoust. Soc. Am.* 126, 2670-2682.
- [5] Kleinschmidt, D. F., Jaeger, T. F. (in press). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychol. Rev.*
- [6] Pavlick, E., Post, M., Irvine, A., Kachaev, D., Callison-Burch, C. (2014). The language demographics of Amazon Mechanical Turk. *Trans. Assoc. Comp. Ling.* 2, 79-92.
- [7] Bernard, A. (2015). An onset is an onset: Evidence from abstraction of newly-learned phonotactic constraints. *J. Mem. Lang.* 78, 18-32.
- [8] Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *J. Exp. Psychol.: Hum. Percept. Perform.* 25, 1568-1578.
- [9] Bernard, A., Onishi, K. H., Seidl, A. (2014). Learning of sound patterns in adulthood: The role of allophonic status. Ms., McGill University, Purdue University.
- [10] Seidl, A., Cristià, A., Bernard, A., Onishi, K. H. (2009). Allophonic and phonemic contrasts in infants' learning of sound patterns. *Lang. Learn. Dev.* 5, 191-202.
- [11] Barr, D. J., Levy, R., Scheepers, C., Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* 68, 255-278.
- [12] Schumacher, R.A, Pierrehumbert J.B., LaShell, P. (2014) Reconciling inconsistency in encoded morphological distinctions in an artificial language. In Bello, P., Guarini, M., McShane, M., Scassellati, B. (eds.) *Proc. 36th Cogn. Sci. Soc.* Austin, TX: Cognitive Science Society, 2895-2900.