Cross-linguistic differences in the perception of palatalization

Molly Babel and Keith Johnson

University of California, Berkeley
mbabel@berkeley.edu, keithjohnson@berkeley.edu

ABSTRACT

This paper investigates the difference between basic psycho-acoustic auditory perception and language-specific perception of speech sounds. This was examined in two experiments with American English and Russian listeners. Results suggest that listeners' language does not influence auditory perception, but does affect the rated perceptual similarity of speech sounds.

Keywords: speech perception, palatalization, English, Russian.

1. INTRODUCTION

Language experience clearly affects the processing of speech sounds [1], but the extent to which language experience affects all speech processing is not well understood. In a speeded AX-discrimination task with fricatives, Dutch and American English listeners patterned identically; language background did not affect their responses. In a similarity rating task, Dutch and American English listeners demonstrated linguistic knowledge and rated fricative pairs differently [2].

A series of experiments with native speakers of Russian and American English Russian language learners reported in [3] found striking effects of language experience on the perception of palatalized consonant sequences. In a labeling task Russian listeners performed *worse* than Russian language learners in identifying C^{j}jV and C^{j}ijV sequences. C^{j}ijV is a rare sequence and carries a low functional load in Russian. C^{j}jV and C^{j}ijV sequences are considered a near-merger; while speakers of Russian produced C^{j}jV and C^{j}ijV sequences differently, they are not sensitive to the acoustic differences. Adult language learners with tentative knowledge of the Russian phonological system hone in on the acoustic differences and perceive the system phonetically, and not linguistically. Interestingly, advanced Russian language learners performed worse on the labeling task than novice language learners, regarding the low functional load contrast between C^{j}jV and C^{j}ijV; that is, they responded to the stimuli more like native speakers, demonstrating their "improved functional knowledge" of Russian.

The two experiments described in this paper investigate the effect of language on similar palatalized consonant sequences with native Russian listeners and American English listeners who have not had *any* experience with Russian.

2. EXPERIMENT 1

The purpose of Experiment 1 is to understand the underlying psycho-acoustic perception of palatalization. It is predicted that language background will not influence the results of this experiment as listeners will respond to acoustic properties of sounds alone.

2.1. Subjects

Fifteen native speakers of American English and fourteen native speakers of standard Russian participated as listeners in Experiment 1. Subjects reported no speech, language, or hearing disorders. Participants were compensated $10 for their time. The data collected from one American English listener and three Russian listeners were excluded from the analysis due to excessively slow reaction times (mean reaction time > 700 ms).

2.2. Stimuli

The stimuli were open syllables with onsets. Possible onsets were: /m/, /v/, /b/, /d/, /l/, and /r/. Consonants were produced with varying degrees of palatalization, and then followed by a vowel: /a/, /u/, or /ɨ/. The degrees of palatalization can be divided into four levels; so, for example, a set of stimuli from a single consonant and vowel combination was *da*, *d^{j}a*, *d^{j}ja*, and *d^{j}ija*. Throughout the course of this paper these types of syllables are referred to as CV, C^{j}V, C^{j}jV, and C^{j}ijV, respectively. A female native speaker of Russian produced the syllables (6 x 3 x 4 = 72) in a carrier phrase.

2.3. Procedure

These naturally produced, isolated CV syllables were presented in pairs with a 100 ms ISI binaurally to listeners over headphones at a workstation using E-prime Experiment Software. Listeners were instructed to judge whether the two tokens were *same* or *different* and log responses on a button box. Listeners were encouraged to keep their reaction times under 500 ms. Reaction time feedback was presented on the computer screen, so listeners could monitor their response time and keep it near 500 ms. Stimuli were blocked by consonant and vowel for a total of 21 blocks. In each block the four "same" pairs were presented three times each and the six "different" pairs were presented twice. This creates a total of 432 pairs throughout the course of the experiment.

2.4. Results and Analysis

The percent correct by language for each different stimulus pair is shown in Table 1. Paired *t*-tests revealed that only three pairs were significantly more difficult for American English listeners.

Table 1: Percent correct identification of pairs by Russian (n=11) and English (n=14) listeners averaged across vowels and consonants. Pairs marked with * are those in which the two language groups differed significantly in identification accuracy at the $p < 0.05$ level.

pair	*Russian*	*English*	*pair*	*Russian*	*English*
bʲijV bʲjV	79	76	mʲijV mʲjV	70	59
bʲijV bV	98	96	mʲijV mV	98	99
bʲV bʲijV	89	89	mʲV mʲijV	92	96
bʲV bʲjV	94	91	mʲV mʲjV	94	96
bʲjV bV	94	94	mʲjV mV	98	99
bʲV bV	98	97	mʲV mV	99	98
*dʲijV dʲjV	93	78	rʲijV rʲjV	82	77
dʲijV dV	99	96	rʲijV rV	100	99
*dʲV dʲijV	96	92	rʲV rʲijV	99	95
dʲV dʲjV	95	98	rʲV rʲjV	92	89
dʲjV dV	98	97	rʲjV rV	98	97
dʲV dV	98	97	rʲV rV	96	98
lʲijV lʲjV	93	85	vʲijV vʲjV	94	87
lʲijV lV	98	100	vʲijV vV	99	97
lʲV lʲijV	92	95	vʲV vʲijV	94	95
lʲV lʲjV	98	98	vʲV vʲjV	90	89
*lʲjV lV	100	94	vʲjV vV	94	97
lʲV lV	100	98	vʲV vV	98	95

"Same" pairs of stimuli and incorrect responses were removed prior to the reaction time analysis. Log reaction time was entered as the dependent factor into a repeated measures ANOVA, and vowel context, listener language, and degree of palatalization were added as independent variables. The analysis revealed a main effect for degree of palatalization (F[5, 91] = 157.69, $p < 0.001$) and vowel (F[2, 26]=7.73, $p < 0.01$). Tukey's HSD shows CʲijV/CʲjV pairs were labeled "different" more slowly than all other pairs, responses to CʲijV/CʲV and CʲV/CʲjV were slower than to CʲijV/CV, CʲjV/CV, and CʲV/CV.

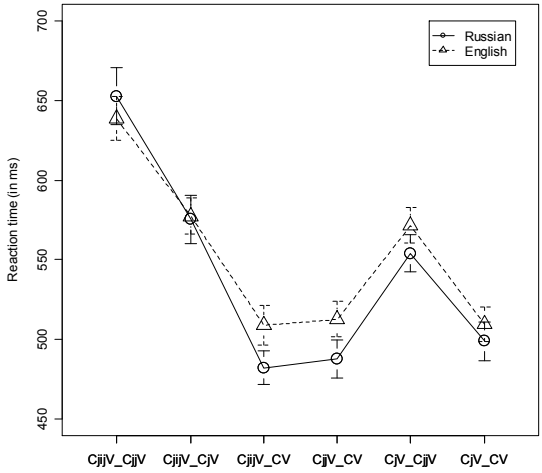

Figure 1: Mean reaction time to pairs of segments palatalized to varying degrees in a speeded AX discrimination task.

There was a significant interaction between degree of palatalization and vowel (F[10, 230]=6.83, $p < 0.001$). Post-hoc tests with Tukey's HSD show that reaction time in the context of /i/ is significantly slower than both /a/ (p <0.001) and /u/ (p <0.05) and reaction times with /u/ were also significantly slower than those with /a/ ($p < 0.05$). Listeners' reaction times in the varying vowel environments are plotted in Figure 2.

2.5. Discussion

Our prediction was borne out in this experiment: language background did *not* affect listeners' responses to the acoustic properties of the stimuli. Regardless of language background, listeners' processing of the stimuli was facilitated when one member of the pair did not have any palatalization. That is, having a simple CV token in a stimulus pair significantly reduced reaction time. Vowel environment also played a significant role. Pairs with /a/ were processed quickest, followed by

those with /u/. The most challenging vowel environment was /i/.

Though these vowel and consonant effects are phonetically interesting, we are primarily interested in whether the low functional load of the $C^{j}ijV$ and $C^{j}jV$ contrast would result in lower perceptual separation of these for Russian listeners as compared with listeners who do not speak Russian. Thus, interactions involving the pair and language factors were of primary theoretical interest for us. Experiment 2 explores the influence linguistic experience has on the perception of speech sounds.

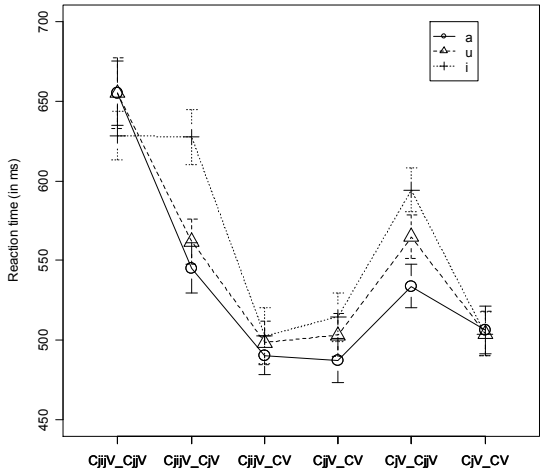

Figure 2: Mean reaction time by vowels in different palatalization contexts in a speeded AX discrimination task. [i] in the figure is /i/.

3. EXPERIMENT 2

The purpose of Experiment 2 is to examine the subjective language-specific organization of speech sounds by asking subjects to rate the perceptual similarity of sounds. It is predicted that language will strongly influence the rated similarity of the sounds.

3.1. Subjects

Thirteen native speakers of American English and ten native speakers of standard Russian participated as listeners in Experiment 2. Subjects reported no speech, language, or hearing disorders. Participants were compensated $10 for their time.

3.2. Stimuli

The same stimuli were used in Experiment 1. Different pairs were presented twice while identical pairs were presented once for each vowel and consonant combination for a total of 288 pairs.

3.3. Procedure

Stimuli were presented in pairs with a 100 ms ISI binaurally to listeners over headphones at a workstation using E-prime Experiment Software. Listeners were instructed to rate the similarity between the two tokens on a 5-point scale. Participants logged their responses on a 5-point equal-interval button box where the buttons had the following labels: [1] *very similar*, [2] *somewhat similar*, [3] *moderately different,* [4] *somewhat different*, and [5] *very different*. Listeners had up to five seconds to respond before the presentation of the next set of stimuli.

3.4. Results and Analysis

An ANOVA with listener rating responses from the different pairs as the independent variable and vowel context, listener language, and degree of palatalization as the dependent variables found main effects for language ($F[1, 21]=9$, $p < 0.01$), degree of palatalization ($F[5, 105]=369.2$, $p < 0.001$) and vowel ($F[2, 42]=21$, $p < 0.001$).

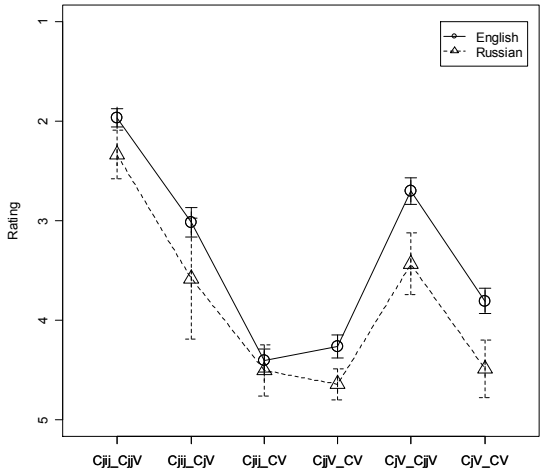

Figure 3: Mean rated perceived similarity of stimulus pairs. A rating of [1] is very similar and [5] very different.

There were significant interactions between degree of palatalization and vowel ($F[10, 210]=18$, $p <$

0.001) and palatalization and language (F[5, 105]=3.4, $p < 0.01$).

Figure 3 shows the rated similarity of the pairs by language. The palatalization by language interaction indicates that Russians and naïve American English listeners perceive the degrees of palatalization somewhat differently. Post-hoc Tukey tests showed significant differences in language groups' rating of all pairs except $C^{j}ijV/C^{j}jV$ and $C^{j}ijV/CV$.

Post-hoc tests with the vowels show pairs in the context of /a/ were rated more dissimilar from those with /u/ or /i/ ($p < 0.001$). Rating averages were also significantly different between /u/ and /i/ ($p < 0.05$); pairs with /u/ were rated more different sounding than those with /i/. The rated similarity of the palatalized pairs by vowel is presented in Figure 4.

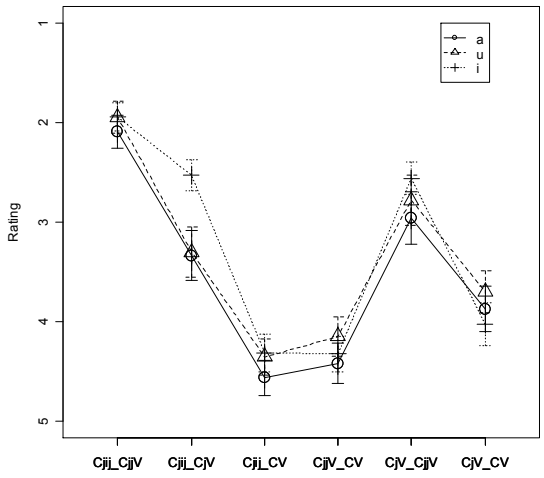

Figure 4: Mean rated perceived similarity of pairs by vowel context. A rating of [1] is very similar and [5] very different.

3.5. Discussion

Post-hoc analyses confirmed our prediction that language experience affects the perceptual rated similarity of speech sounds. Native speakers hear greater contrast among degrees of palatalization; Russian listeners rated the sounds to be more different than American English listeners across most of the pairs. However, for two of the three pairs involving the functionally rare sequence $C^{j}ijV$ Russians' ratings were not different from Americans'. The native speaker enhancement was not found.

4. CONCLUSION

This set of experiments illustrates that in a behavioral task, Russian and naive American English listeners perceive speech sounds the same on a psycho-acoustic level, but slightly differently in terms of subjective perceptual organization of the same sounds. The patterns of Figures 1 and 3 demonstrate how pairs that are acoustically similar (= slower reaction times) are the same as those that are rated as more similar sounding for both Russian and American English listeners.

Russian speakers produce $C^{j}ijV$ and $C^{j}jV$ sequences differently and, in this experiment, they are equally sensitive to the perceived differences of these sounds as naïve American English listeners. The perceived similarity of $C^{j}ijV$ and $C^{j}jV$ are most prominent when they are presented as a pair or when $C^{j}ijV$ is paired with CV. The native speaker tendency to rate palatalization contrasts as more different than non-native speakers is not found with most pairs involving the functionally rate sequence $C^{j}ijV$. Across languages, these pairs of sequences are perceived to be equally similar sounding. Adult Russian language learners have been found to be more sensitive to these cues than native speakers of Russian, but it appears that American English listeners with *no* Russian language experience perceive these sounds in ways similar to Russian speakers.

5. ACKNOWLEDGMENTS

This project has been funded by NIH grant #R01 DC004421 awarded to the second author. We would like to thank all of our research subjects for their participation.

6. REFERENCES

[1] Näätänen, R., Lehtoski, A., Lennes, M. Cheour, M., Houtilainen, M., Iivonev, A., Vainio, M., Alku, P. Iimoniemi, R., Luuk, A., Allik, J., Sinkkonen, J, Alho, K. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.

[2] Johnson, K and Babel, M. in prep. Perception of fricatives by Dutch and English listeners.

[3] Diehm, E. 1998. *Gestures and linguistic function in learning Russian: Production and perception studies of Russian palatalized consonants*. Ph. D. Dissertation, The Ohio State University.