Utilizing electropalatography to train palatalized versus unpalatalized consonant productions by native speakers of American English learning Russian

Jane F. Hacking, Bruce L. Smith and Eric M. Johnson
University of Utah

Previous research has shown that English-speaking learners of Russian, even those with advanced proficiency, often have not acquired the contrast between palatalized and unpalatalized consonants, which is a central feature of the Russian consonant system. The present study examined whether training utilizing electropalatography (EPG) could help a group of Russian learners achieve more native-like productions of this contrast. Although not all subjects showed significant improvements, on average, the Russian learners showed an increase from pre- to post-training in the second formant frequency of vowels preceding palatalized consonants, thus enhancing their contrast between palatalized and unpalatalized consonants. To determine whether these acoustic differences were associated with increased identification accuracy, three native Russian speakers listened to all pre- and post-training productions. A modest increase in identification accuracy was observed. These results suggest that even short-term EPG training can be an effective intervention with adult L2 learners.

Keywords: electropalatography, Russian, consonants, palatalization, training study

1. Introduction

The distinction between unpalatalized and palatalized consonants is a central feature of the Russian consonant system. For example, 'mother' (*/matʲ/*) and 'checkmate' (*/mat/*) are differentiated by their final consonant: palatalized /tʲ/ for 'mother' and unpalatalized /t/ for 'checkmate'. Articulatory descriptions of Russian palatalization describe it as a secondary articulation in which the tongue is raised up and forward toward the hard palate (Avanesov, 1972; Ordin, 2010). Acoustic

Journal of Second Language Pronunciation 3:1 (2017), **9–33**. DOI 10.1075/jslp.3.1.01hac
ISSN 2215–1931 / E-ISSN 2215–194X

analysis has shown that Russian palatalized consonants are phonetically distinguished from their unpalatalized counterparts by two main acoustic cues: formant transitions of preceding vowels and characteristics of the consonant release bursts (Bolanos, 2013; Kochetov, 2006). Vowels before palatalized consonants have final F2 values that are 300–500 Hz higher than those before unpalatalized consonants (Fant, 1970; Purcell, 1979). Release bursts of palatalized consonants are longer in duration than those of unpalatalized consonants (Kochetov, 2006; Ordin, 2010). The twelve unpalatalized/palatalized consonantal pairs of Russian are shown in Table 1.

Table 1. Russian consonants paired for the feature of palatalization

	Labial		**Dental/Alveolar**	
	unpalatalized	**palatalized**	**unpalatalized**	**palatalized**
Voiceless stop	p	p^j	t	t^j
Voiced stop	b	b^j	d	d^j
Voiceless fricative	f	f^j	s	s^j
Voiced fricative	v	v^j	z	z^j
Nasal	m	m^j	n	n^j
Lateral liquid			l	l^j
Trill			r	r^j

Previous research has demonstrated the difficulty English-speaking learners of Russian have acquiring the palatalization contrast. Diehm (1998) reported acoustic data from a production study in which native speakers and L2 learners were recorded reading lists of words containing various consonant targets in three environments: /C^jV/, /C^jjV/ and /C^jijV/. She concluded that L2 Russian speakers produce palatalized sequences with closures that are too long for /C^jV/ and too short for /C^jjV/. Hacking (2011) showed that Russian native-speaker listeners could not reliably differentiate between advanced proficiency L2 learners' productions of unpalatalized and palatalized consonants in word-final position. In a forced choice word identification task, Russian native listeners identified learners' productions of both /kon/ and /kon^j/ as /kon/, suggesting an inability on the part of the L2 speakers to produce the contrast. Hacking, Smith, Nissen, and Allen (2016) provided acoustic and electropalatographic (EPG) measurements of native and learner productions of four unpalatalized/palatalized pairs (/t/-/t^j/, /p/-/p^j/, /l/-/l^j/, and /s/-/s^j/). These data showed clear acoustic and EPG differences between unpalatalized and palatalized consonants for native speakers' productions, whereas no differences were evident among learners' productions.

The learners in these three studies all had significant experience studying Russian and living in Russian speaking environments. Through classroom instruction using standard materials, they had been introduced to articulatory descriptions of palatalization and had opportunities to practice the contrast. In addition, the extended time spent living or studying in a Russian speaking country ensured that all subjects had extensive native speaker input. The documented difficulties that even experienced Russian learners have with palatalized consonants raise the question whether there is a more effective way to help learners acquire this important phonological contrast.

2. L2 training studies

In recent years technological advances have provided various methods for attempting to enhance foreign language instruction, including L2 pronunciation teaching. For example, Ruellot (2011) used spectrograms to help intermediate-level English-speaking learners of French improve their pronunciation of the French phonemes /u/ and /y/. Patten and Edmonds (2015) used spectrographic feedback with Japanese learners of English to improve their pronunciation of English /ɹ/. Kartushina, Hervais-Adelman, Frauenfelder, and Golestani (2015) worked with learners of Danish to improve productions of four vowels by providing learners with visual plots of F1 and F2 native vowel targets and comparison plots of their own productions during training. Other researchers have experimented with Automatic Speech Recognition training (Demenko, Wagner, & Cylwik, 2010; Machovikov, Stolyarov, Chernov, Sinclair, & Machovikova, 2002) and/or visual feedback mechanisms (Engwall, 2012; Massaro, Liu, Chen, & Perfetti, 2006) as methods for L2 pronunciation improvement.

Another technique that has been used in L2 pronunciation training is electropalatography (EPG). Although EPG has been used by a number of speech pathologists to treat various types of articulation deficits (e.g., Cleland, Timmins, Wood, Hardcastle, & Wishart, 2009; Fujiwara, 2007; Gibbon & Wood, 2003; Nordberg, Carlsson, & Lohmander, 2011), relatively few studies have utilized EPG to assist in training L2 learners. Gibbon, Hardcastle, and Suzuki (1991) trained two Japanese learners of English on the /ɹ/-/l/ distinction. Both subjects were able to produce the distinction in controlled contexts after four 45-minute training sessions. Schmidt and Beamer (1998) reported data from a training study with three native speakers of Thai that targeted three English contrasts: /θ/-/t/, /tʃ/-/ʃ/, and /ɹ/-/l/. Results showed that all three improved their articulatory positioning, and two subjects could consistently produce the /ɹ/-/l/ distinction (in controlled contexts). Schmidt (2012) also explored links between perception and production of the English contrasts /s/-/ʃ/, /z/-/dʒ/, and /ɹ/-/l/ by two adult native speakers of Korean who participated in EPG training. Subjects learned to produce the contrasts, and perception improved for trained, but not for untrained consonants. Because EPG can provide visual feedback of tongue palate contact, which is a critical aspect of the Russian palatalization constrast, the present training study employed this approach to examine whether L2 Russian learners could improve their production of the contrast between palatalized and unpalatalized consonants.

The training protocol described in section 3.2 was based on principles outlined in the tutorial, *Principles of Motor Learning in Treatment of Motor Speech Disorders* (Maas et al., 2008). The tutorial reviews studies of motor learning and investigates their applicability to the treatment of motor speech learning. Of particular relevance is the review of various aspects of treatment design: length of treatment, frequency of treatment sessions, blocked versus random sessions, and type of feedback. The authors concluded that there is no empirical evidence regarding an optimal amount of practice for speech motor learning. In addition, there does not appear to be any advantage to spreading the same number of training sessions over a shorter or longer time period (p. 283). Treatment design can present different movements in a blocked (subjects work on a single target repeatedly before moving on to a different one) or random fashion (two or more targets are presented randomly). They reference one study comparing blocked and random practice with unimpaired speakers (Adams & Page, 2000), which found that random practice produced more effective results. They also presented studies comparing two types of feedback: knowledge of results (KR) and knowledge of performance (KP), concluding that both types of feedback are generally equally effective but that KP feedback may be particularly helpful when the movement is novel or components of it are unclear. KR is when particpants receive information about movement outcome in relation to the goal. In speech therapy settings, this is often realized as specific commentary by the therapist regarding whether a particular target was achieved. KP refers to feedback that characterizes the nature or quality of the movement pattern. Biofeedback falls under this category of feedback, as does commentary by a therapist that tells the participant how to alter perfomance to achieve a better result.

3. Methods

3.1 Participants

Ten students who were majoring or minoring in Russian and were enrolled in an upper division university course on Russian phonetics and phonology during Spring semester 2014 served as subjects. All 10 students participated in an

EPG training protocol as part of the course curriculum. In terms of their levels of Russian language expertise, two students were enrolled in fourth semester Russian, three students were enrolled in sixth semester Russian, four students had learned Russian while serving a 22-month religious service mission in Russia (RM – Returned Missionary), and one student's primary language training was a 48 week intensive course (7–8 hours per day, 5 days a week) at the Defense Language Institute. All but one were native speakers of English. This student's native language was Serbian, which like Russian is part of the Slavic language family, but crucially does not have the palatalization contrast found in Russian. None of the participants reported any speech, language, hearing, or neurological disorders. The project was approved by the university's Institutional Review Board, and all subjects signed consent forms explaining the nature of the investigation and informing them of their right to withdraw from participation at any time.

Students were grouped by language background, and then members in each of these language background sub-groups were randomly assigned to either Group 1 or Group 2. Thus each group was balanced in terms of Russian language background. Table 2 provides details about the 10 participants and the two groups they were assigned to. Because the training was part of the curriuclum for the Russian phonetics and phonology course in which the students were enrolled (and consistent with a majority of EPG training studies, e.g., Lohmander, Henriksson, & Havstam, 2010; McAuliffe & Cornwell, 2008), no control group was included, i.e., there would have been no opportunity to provide subsequent comparable training for control subjects.

Table 2. Participant background information

ID #	Gender	Language Background	Group
1	M	RM	1
2	M	RM	1
3	F	2nd year	1
5	F	3rd year (native language = Serbian)	1
6	F	3rd year	1
4	M	RM	2
7	M	Completed Defense Language Institute	2
8	F	2nd year	2
9	M	3rd year	2
10	M	RM	2

3.2 Materials and procedures: EPG training and recordings

At the beginning of the course, students visited the university dental clinic, where a mold was taken of their palate. These molds were sent to the company Complete Speech, where a ‘pseudopalate’ was made for each student. Each ‘pseudopalate’ contained 124 electrodes arranged as pictured in Figure 1.

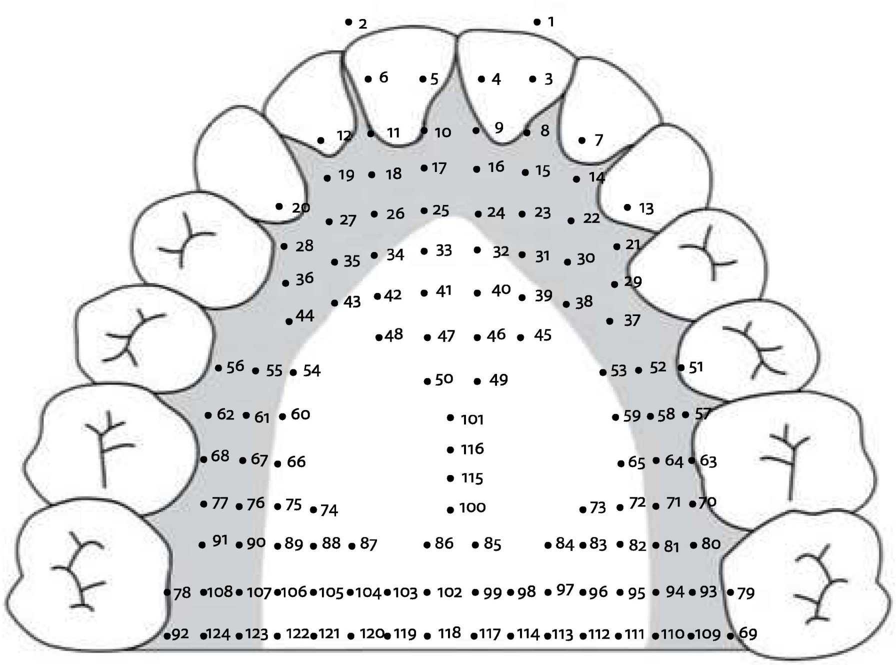

Figure 1. Diagram showing the complete speech pattern of the 124-electrode array worn during training

Materials for this study included a script of Russian sentences with words containing 12 target consonants, EPG visual display targets for the trained consonant pairs, and listening files of native speaker productions of the target consonants. Words in the script contained the sounds /p/-/pʲ/, /t/-/tʲ/, /s/-/sʲ/, /n/-/nʲ/, /l/-/lʲ/, /r/-/rʲ/ in word- and sentence-final position. The decision to focus on coda position was motivated by prior research documenting the tendency of English speakers to realize a sequence of palatalized consonant + vowel as an unpalatalized consonant + glide + vowel (Antonova 1988; Bryzgunova 1963). The coda environment isolates the contrast under investigation, providing the clearest evidence of how articulatory behavior differs with palatalization for native speakers, thus allowing us to investigate what patterns non-native speakers exhibit when attempting to differentiate palatalized from unpalatalized consonant targets. Of these, a subset was trained (/t/-/t^j/, /s/-/s^j/) while the rest were untrained. The 12 target words each occurred five times throughout the script. Words were presented in the carrier phrase: *Vot opjat' slovo* _________/ '*Here again is the word*' __________. Each page of the script began and ended with two carrier phrases with filler words not containing any of the target segments. Participants were recorded reading the script prior to training, after the first two weeks of training, and after the second two weeks of training. For training, we created visual EPG targets for four consonants: /t/-/t^j/, /s/-/s^j/, along with listening files containing native speaker productions of these same consonants in Russian words and phrases. The visual targets were created from EPG recordings of native speaker productions made using the Smart Palate software developed by Complete Speech. The software allows a user to capture and save a target and make manual adjustments to the visual display. Figure 2 shows screenshots of the four targets. Black dots indicate electrodes; blue circles indicate electrodes activated at the point of maximum contact for a given consonant. During training, students were instructed to attempt to activate electrodes within the blue circles for each of the respective consonants.

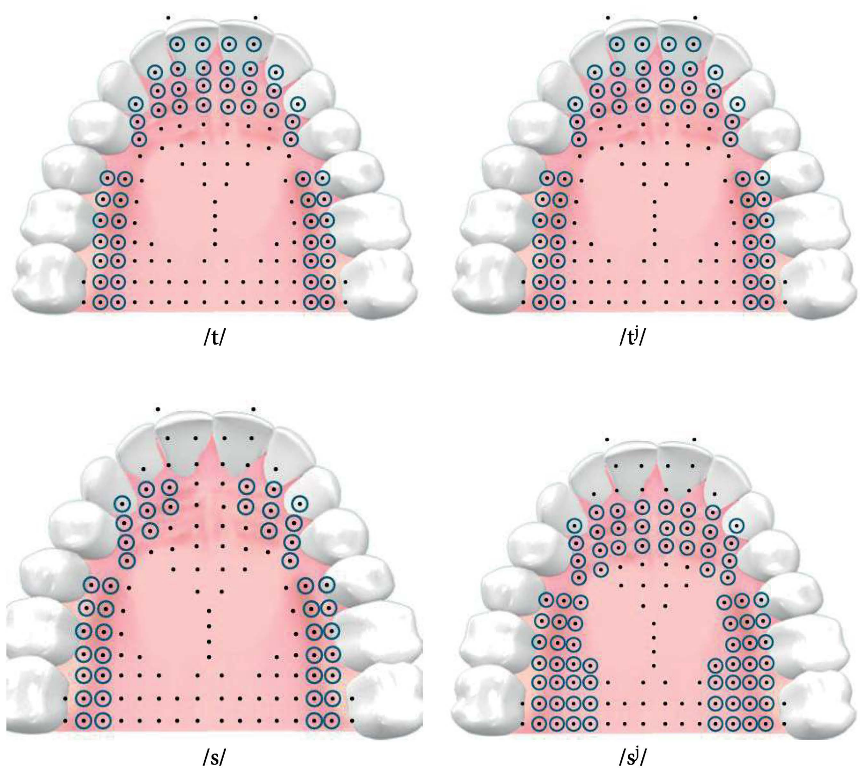

Figure 2. Screenshots of target consonants that students saw during training

A training protocol that students could complete independently was designed. Most EPG training studies reported in the literature have been conducted with heavy therapist/researcher involvement, i.e., each training session is directly supervised (e.g., McAuliffe & Cornwell, 2008). One of our goals was to create conditions that could be easily replicated with students working on their own. Additionally, we were guided by the principles identified in Maas et al. (2008) outlined in section 2 above. Without clear evidence supporting a particular length or frequency of training sessions, we focused on creating a regimen that was easy to implement: two two-week training periods comprising eight short individual sessions per training period. Consistent with evidence that KP feedback can be helpful when the movement is novel, all feedback students received was of the KP variety. In a preparatory session with one of the researchers (Hacking), each student received explicit feedback about tongue placement for the target consonants. During actual training sessions, students monitored their own tongue placement against the representative EPG visual targets created from native speaker productions. While there is limited evidence that random practice is more effective than blocked practice (Adams & Page, 2000; Maas et al., 2008), the constraints of the software (only one target can be displayed on the screen at a time, and the process of changing visual targets is somewhat slow) meant that we provided blocked practice.

Prior to the start of training, one class period (80 minutes) was used to orient the students as a group to their pseudopalates and introduce them to the Smart Palate software interface. Then, the lead researcher met with each student individually for 15–20 minutes. These one-on-one sessions were used to go through the training protocol step by step, ensuring the student could complete the training independently. Students had to demonstrate the ability to open the Smart Palate program, connect their pseudopalate, access the listening files, and load the appropriate EPG tongue targets. The researcher reminded them that the target images were idealized and worked with them to see how their own mouth architecture might result in variations in contact pattern. They were reminded to focus on approximating the two targets and differentiating between them.

Prior to the onset of training, each subject was audio-recorded reading the script described above to obtain a baseline of his/her productions of palatalized and unpalatalized consonants. Because the training was delivered as part of a college course, there was not a control group that received no training. To somewhat offset this limitation, training was divided into two parts and the participants into the two groups previously described. During the first two-week training period, one group worked on the /t/-/t^j/ contrast, and the other group worked on the /s/-/s^j/ contrast. During each of the two training periods, participants completed eight 15-minute training sessions over a two-week period. They were instructed not to complete sessions in immediate succession, but were allowed to do two sessions in a day if they took at least a one hour break. At the end of the first two-week training period, each participant was recorded a second time reading the script. During the second two-week training period, participants switched to train on the other consonant pair. After the second training period, they were recorded a final time. Recordings were done without the pseudopalate in place since our aim was to gather natural data from which to measure the acoustic properties of their productions.

Participants received a detailed protocol for the training sessions. This included a set of instructions reviewing set-up procedures and instructions for each of the eight individual training sessions. Each session began with participants listening for three minutes to segments, words or sentences (depending on the session) containing the target segments. Session 1 was devoted to production of the two segments in isolation. Students were instructed to "focus on the distinction between the tongue/palate contact pattern for the two sounds." For sessions 2 and 3, students were asked to spend the first 1–2 minutes on the individual segments as in session 1 and then for the remainder of the time to move on to a list of syllables with the targeted consonants in final position (e.g., ат – ать /at – at^j/). In session 4, students began by working on the syllables from sessions 2 and 3 and then moved on to words with the unpalatalized consonant in word-final position (e.g., брат /brat/). Session 5 had the same structure as session 4, but with practice of palatalized consonants in word-final position (e.g., брать /brat^j/). Sessions 6 and 7 required practice of the same words as in the previous two sessions but in a single session. Students were instructed to set their visual target for one consonant, go through the list of words and then switch to the other visual target and the next set of words. They repeated this until they had reached the time limit of 15 minutes for the session. The procedure for the eighth and final session was similar to session 7, except that the words were presented in the carrier phrase used in the recording script: *Vot opjat' slovo* _________/ '*Here again is the word*' __________.

3.3 Materials and procedures: Native speaker listening task

To assess the degree to which the learners' productions were perceptible to native-speaker listeners, we conducted a listening task. Three female native speakers of Russian between the ages of 25 and 35, who had been in the United States for between 4 and 10 years, completed the listening task. Tokens from the learners' pre-training and post-training recordings were presented as a forced-choice word identification task requiring the participant to press a left or right key on a keyboard depending on which word on the screen they thought they heard. The five recordings of each item produced by each L2 speaker both pre- and post-training were presented. These 1200 word tokens (12 words × 10 tokens × 10 learners) were presented in six blocks with five subject-controlled breaks between blocks. Tokens were presented randomly, so listeners did not hear all tokens from a single speaker at once, nor did they hear members of minimal pairs in juxtaposition. Stimuli were presented and results obtained using DMDX experiment presentation software (Forster & Forster, 2003). The task took approximately 55 minutes, after which each participant completed a brief questionnaire to gather basic biographical information, details about her study and exposure to English and other languages, and to ascertain whether she had any hearing difficulties.

4. Data analysis

The primary measures of interest were second formant (F2) frequency of the vowel preceding the palatalized and unpalatalized consonants and target consonant release noise duration. Acoustic analyses of the data were performed utilizing Praat 5.3.23 (Boersma & Weenink, 2012), which simultaneously displayed a waveform and spectrogram of each target word. The settings utilized included: Burg method with 50 Hz pre-emphasis, frequency range from 0–5000 Hz, Gaussian window shape, and window length of .005 ms. Vowel duration, consonant closure duration, and consonant release (burst) noise duration were measured for each target word, as were F2 at the vowel midpoint and vowel endpoint, and halfway between those two times (i.e., at .75 vowel duration) for each token. Segmentation of the consonants and vowels was based on commonly-utilized acoustic characteristics associated with substantial changes in waveform shape and/or amplitude, the occurrence of consonant release bursts, and other salient acoustic events related to frequency and amplitude characteristics of the waveform and spectrographic displays (Smith, Hillenbrand, & Ingrisano, 1986).

Intra-judge measurement reliability was assessed by having the same individual who made all the original acoustic measurements re-measure the data for one of the 10 subjects a second time after a period of approximately six months. Inter-judge measurement reliability was assessed by having a second person also measure the data for one of the subjects. Intra-judge reliability averaged 53 Hz (5%) for F2 frequency, and inter-judge reliability was 54 Hz (4%). Intra-judge reliability for consonant release noise was 7 msec (13%), and inter-judge reliability was 13 msec (17%).

The data captured using the DMDX experiment presentation software for the native speaker listening task contained information about each token and whether it was correctly identified by each listener. These data were sorted by speaker, by word, and whether the recording was made pre- or post-training. For each listener, percent correct identification was calculated for each speaker and for each word pre- and post-training.

5. Results

As outlined in section 3.1, we divided participants into two groups and had each group train on one of the two palatalized versus unpalatalized pairs first and switch to the other pair for the second set of training sessions. Group 1 trained first on the /t/-/t^j/ contrast, while Group 2 began with the /s/-/s^j/ contrast. There were no substantive findings related to this division; that is, neither group performed differently. Therefore, in the discussion of results that follows, we report data for the two groups together.

5.1 Acoustic analysis: Group data

One of the primary acoustic characteristics marking the distinction between Russian palatalized and unpalatalized consonants is the F2 of the vowel preceding the target consonant (Bolanos, 2013; Kochetov, 2006). Figure 3 shows average F2 values at the conclusion of the vowel for the group of 10 subjects averaged across the six pairs of palatalized versus unpalatalized target consonants for three time periods, viz., pre-training, mid-training, and post-training. A one-way repeated measures ANOVA indicated that there was a significant difference in F2 frequency for vowels preceding unpalatalized versus palatalized consonants across the three time periods [$F\ (5,9) = 22.78$, $p < .0001$]. A Tukey's multiple comparisons test determined that there was a significant difference between the 10 Russian learners' average F2 for vowels preceding unpalatalized consonants (1280 Hz) at pre-training versus their average F2 for vowels preceding palatalized consonant targets (1420 Hz), which is a difference of 140 Hz or 11% ($p < .05$). At post-training, average F2 for vowels preceding unpalatalized consonants was 1279 Hz, whereas average F2 for vowels preceding palatalized consonant targets was 1604 Hz, a difference of 325 Hz or 25% ($p < .01$). The 270 Hz (21%) difference in F2 frequency for vowels preceding palatalized versus unpalatalized consonants at mid-training was also significant ($p < .01$). It is worth noting that although there was already a signficant difference at pre-training in the subjects' F2 values preceding unpalatalized consonants versus palatalized consonants, the contrasts at mid-training and post-training were approximately twice as large.

Figure 3 also shows that the pre- to post-training change for F2 of vowels preceding palatalized consonants was 184 Hz (an increase of 13%, $p < .001$), whereas F2 of vowels preceding unpalatalized consonants remained unchanged from pre- to post-training (i.e., the difference was −1 Hz [0% change], ns). The increase in F2 of 147 Hz from pre-training to mid-training for vowels preceding palatalized consonants was also significant ($p < .01$), whereas the increase from mid-training to post-training (37 Hz) was not. We also note in Figure 3 that average F2 for vowels

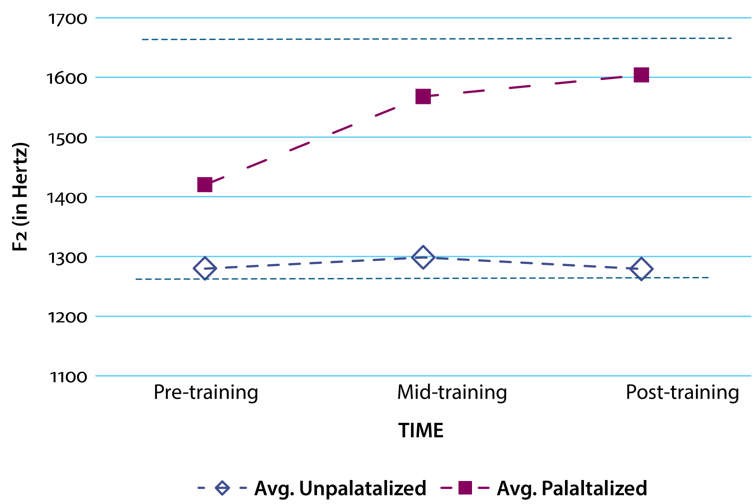

Figure 3. A comparison of vowel F2 preceding palatalized versus unpalatalized consonant targets at pre-, mid-, and post-training. The upper and lower gray dashed lines, respectively, show average vowel F2 values produced by native Russian speakers preceding palatalized versus unpalatalized consonants (Hacking et al., 2016)

preceding unpalatalized consonants produced by six native Russian speakers was 1282 Hz (from Hacking et al., 2016), indicated by the lower dashed gray line, which was identical to the average for the Russian learners' average F2 (1282 Hz) at pre-training. By comparison, the native Russian speakers' average F2 for vowels preceding palatalized consonants (i.e., the upper gray dashed line) was 1675 Hz, which is 255 Hz (18%) greater than the L2 learners' average for palatalized target consonants at pre-training, i.e., 1420 Hz (Mann-Whitney U = 9, $p < .05$). Although the L2 learners' post-training average F2 for vowels preceding palatalized targets (1604 Hz) was still less than the native Russian speakers' F2 average value (1675 Hz, a difference of 4%), the L2 learners' average fell within the range shown across the six L1 Russian speakers (1547–1818 Hz), and their group averages were not significantly different from one another (Mann-Whitney U = 25, ns).

As indicated previously, because the subjects were students in a one-semester Russian phonetics and phonology course and the training was part of the course curriculum, it was not appropriate to select a subset of students to serve as a control group that would not receive any training on the palatalization contrast. As an alternative we examined how subjects performed on trained and untrained segments. As shown in Figure 4, F2 for vowels preceding trained palatalized consonants (solid line with filled circles) increased by 275 Hz (18%) from pre-training to post-training, and F2 for vowels preceding untrained palatalized consonants (solid line with open circles) increased by 139 Hz (10%). It should be noted that the mid-point results are based on the average values from two sub-groups, i.e., the five subjects that were initially trained on the /t/-/tʲ/ contrast and the five subjects that were initially trained on /s/-/sʲ/ contrast.

When comparing vowel F2 frequencies for the trained and untrained palatalized and unpalatalized consonant targets, a one-way repeated measures ANOVA was significant [$F(9,11) = 26.89$, $p < .0001$]. Tukey's multiple comparison tests indicated that significant changes occurred from pre-training to post-training for F2 of vowels preceding the trained and the untrained palatalized targets ($p < .01$ in both cases), and from pre-training to mid-training for vowels preceding the untrained, palatalized targets ($p < .05$). No significant changes occurred for F2 of vowels preceding the unpalatalized targets (dashed lines with squares) across any of the time periods.

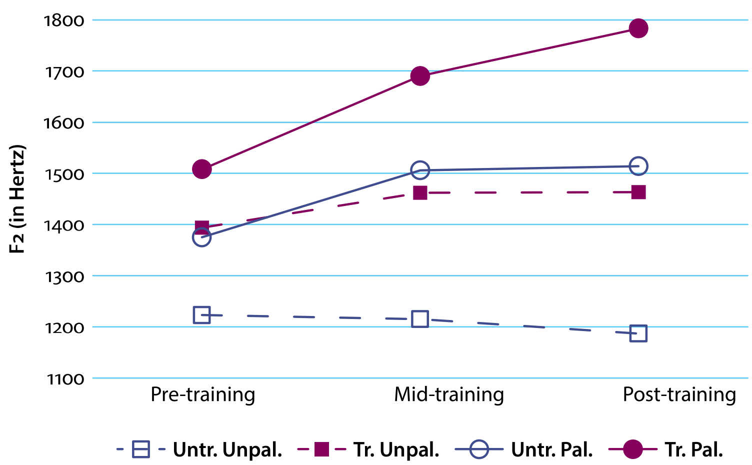

Figure 4. Second formant frequency values for vowels preceding trained versus untrained palatalized and unpalatalized consonant targets at pre-, mid-, and post-training

In addition to looking at the subjects' productions of vowels preceding palatalized versus unpalatalized consonants, we also examined how the L2 learners performed on the individual target words. Table 3 shows average "difference values" (i.e., post-training minus pre-training) for F2 for the 10 Russian learners as a group at the endpoint of vowels preceding the six unpalatalized versus the six palatalized consonant target words. The average lack of change (−1 Hz) in F2 for the unpalatalized consonants that was seen in Figure 3 is reflected in the top row of Table 3 with the vowel F2 difference values being approximately the same (within ± 75 Hz) for all the target words at post-training relative to pre-training. By comparison, the increase in F2 difference values observed in Figures 3 and 4 from pre-training to post-training for vowels preceding palatalized target consonants can be seen in the second row of Table 3 for five of the six individual target words;

/udar^j/ (2 Hz) was the only word that did not show a substantial increase in the F2 difference value (ranging from 114 Hz for /top^j/ to 373 Hz for /atbros^j/) from pre-training to post-training.

Table 3. Average "difference value" (in Hz) for 10 subjects' vowel F2 at post-training (Time 3) versus pre-training (Time 1) preceding each of the six unpalatalized versus the six palatalized target words (* = $p < .05$; ** = $p < .01$). A "+" indicates words that were specifically trained

	/brat/$^+$	**/top/**	**/ugel/**	**/atbros/$^+$**	**/kon/**	**/udar/**	**Avg.**
Post-training minus Pre-training (Hz)	62	1	−58	75	−38	−49	−1
	/brat^j/$^+$	**/top^j/**	**/ugel^j/**	**/atbros^j/$^+$**	**/kon^j/**	**/udar^j/**	**Avg.**
Post-training minus Pre-training (Hz)	176	114	211	373	231	2	185
Difference	114*	113**	269**	298**	269**	51	184

A two-tailed Wilcoxon matched-pairs signed rank test determined that there was a significant difference between the difference values for post-training relative to pre-training ($p < .05$). One-tailed Wilcoxon matched-pairs signed rank tests also indicated that the increase in the difference values from pre-training to post-training for five of the six individual word pairs were statistically significant ($p < .05$ or better, see table), with the only exception being /udar/ versus /udar^j/.

Previous research has indicated that the duration of consonant "release noise" is also an acoustic parameter that distinguishes palatalized from unpalatalized consonants in the speech of native Russian speakers (Bolanos, 2013; Kochetov, 2006), with palatalized consonants being characterized by longer consonant release noise. Therefore, we examined the speech of the 10 Russian learners for this factor. As shown in Figure 5, post-consonant release noise for the L2 group's productions of palatalized consonants increased by an average of 26 msec from pre-training (85 msec) to post-training (111 msec). A Wilcoxon matched-pairs signed rank test indicated that this difference was significant (W = 41.0, $p < .05$). This raises the question as to what these 10 L2 subjects' duration of post-release noise was for unpalatalized target consonants before and after training. Figure 5 shows that, on average, the learners' release noise duration was comparable before (90 msec) and after (85 msec) the training procedure (W = −16.0, ns). These values for unpalatalized release noise are thus comparable to the group's pre-training release noise duration for palatalized target consonants. By comparison, the native Russian speakers' consonant release noise reported in Hacking et al. (2016) averaged 69 msec for unpalatalized consonants and 100 msec for palatalized consonants. Thus, prior to the training protocol, the 10 Russian learners tended to produce post-consonant release noise that was approximately the same in duration for both palatalized and unpalatalized targets (i.e., 85–90 msec), and which was approximately intermediate within the range of the unpalatalized (69 msec) versus palatalized (100 msec)release burst noise by native speakers.

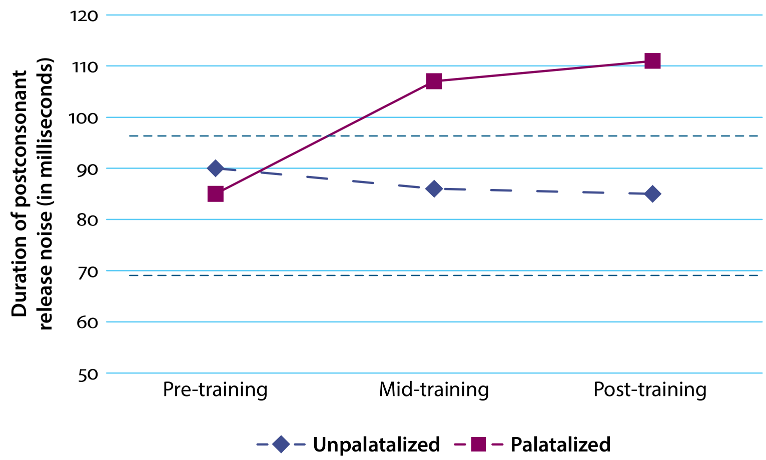

Figure 5. Post-consonant release noise for palatalized versus unpalatalized targets produced by 10 L2 Russian learners. The horizontal dashed lines show the range of values for native Russian speakers (Hacking et al., 2016)

5.2 Acoustic analysis: Individual data

Given that this was a training study, it was reasonable to expect that there would be some differences in the performance of the 10 Russian learners at pre-, mid- and/or post-training. Table 4 shows the findings for each of the 10 Russian learners (i.e., S1–S10) in terms of their "difference value" between F2 of vowels preceding palatalized minus unpalatalized consonant targets for pre-training (row 1), mid-training (row 2) and post-training (row 3), as well as the difference value increase from pre-to post-training (row 4). Each of the 10 subjects showed an increase in their F2 difference value for vowels preceding palatalized versus unpalatalized consonants from pre-training (average group difference value = 140 Hz) to mid-training (average group difference value = 292 Hz), i.e., an increase of 152 Hz (109%). Each subject also showed an increase in their difference value from pre-training (140 Hz) to post-training (average group difference value = 325 Hz), i.e., an increase of 185 Hz (132%). However, not all the subjects showed an increase from mid-training to post-training (i.e., difference values were smaller at post-training relative to mid-training for S4, S5, S6, and S7). As can also be seen in Table 4, the individual subjects' vowel F2 difference values preceding palatalized versus unpalatalized targets ranged from 15 Hz (S7) to 295 Hz (S8) at pre-training, compared to values from 132 Hz (S1) to 588 Hz (S8) at mid-training and from 83 Hz (S7) to 634 Hz (S8) at post-training. Based on the averages of each subject's five repetitions of each of the six target words, individual one-tailed Wilcoxon matched-pairs signed rank tests determined that 7 of the 10 Russian learners (i.e., S1, S2, S3, S4, S8, S9 and S10) showed significant ($p < .05$) increases in their difference values when comparing their post-training with their pre-training values.

Table 4. Ten subjects' average "difference values" for F2 (in Hz) preceding palatalized versus unpalatalized consonants and the increase from pre-training to mid-training to post-training (* = $p < .05$)

	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**Avg.**
Pre-training	32	180	262	229	193	57	15	295	52	78	140
Mid-training	132	345	359	492	337	195	141	588	182	145	292
Post-training	226	456	388	456	314	92	83	634	242	356	325
Pre- to Post-Increase	194*	276*	127*	227*	121	34	68	339*	190*	278*	185

Regarding the duration of consonant release burst noise, the L2 group's average increase from pre- to post-training was primarily due to increases by only 3 or 4 of the 10 subjects. Furthermore, compared to the findings from six native Russian speakers (Hacking et al., 2016), even before the training process was initiated, a majority of the 10 Russian learners were already producing post-consonant release noise that was essentially comparable to the native speakers' values, and they did not change from pre- to post-training. Rather, the L2 group's average increase was a result of 3 or 4 of the subjects with shorter release burst noise durations that showed increases from pre- to post-training, which brought their values into the native speaker range.

5.3 Listener judgments

Since a number of statistically-significant acoustic-based changes were observed in the productions of the 10 Russian learners from pre- to post-training, we were also interested in whether the speakers' palatalized versus unpalatalized consonant productions could be differentiated by native speakers of Russian. To examine whether the palatalization contrast was detectable at pre- and/or post-training, three native Russian speakers completed a forced word choice identification task as described above in section 3.3. As shown in Table 5, the three native Russian listeners, on average, correctly identified 60% of the pre-training tokens produced by the 10 Russian learners versus 68% accuracy at post-training. A paired t-test indicated that this increase was statistically significant (t = 3.480, df = 9, $p < .01$). As can be seen in the table, the native listeners showed the greatest increases in identification accuracy for Subjects 1, 7, and 10, as compared to little or no improvement for Subjects 2, 3, 4, 5, 6 and 8. The results for Subject 9 were somewhat intermediate, not falling clearly into either sub-group. It is also worth noting in Table 5, however, that the subjects who showed the greatest improvement (1, 7, 10) from pre- to post-training were not the ones who were identified most accurately, and the subjects whose productions were identified most accurately (3, 4, 5, 8) were not the ones who showed the greatest improvement.

Table 5. Percent correct identification by three native Russian listeners for 10 Russian learners averaged across all palatalized and unpalatalized target words at pre-training versus post-training

Subject	1	2	3	4	5	6	7	8	9	10	Avg.
Pre-training	56	66	68	72	68	53	44	69	55	51	60
Post-training	68	69	71	70	73	57	64	73	64	66	68
Difference	12	3	3	−2	5	4	20	4	9	15	8

When examining accuracy for the individual words (Table 6), the native listeners were most accurate in identifying the unpalatalized targets /top/ (89%), /kon/ (87%), and /ugel/ (76%) at pre-training, with little change for these three words at post-training. The greatest increases in identification accuracy from pre-training to post-training were for the palatalized counterparts of these words, viz., /top^j/ (increasing from 39% to 65%), /kon^j/ (increasing from 65% to 87%), and /ugel^j/ (increasing from 47% to 62%). Aside from /brat^j/ (which increased from 57% to 71% = 14%), the remaining words showed reasonably small or no increases in identification accuracy from pre-training to post-training – despite the fact that, as seen in Table 3, for example, some of the acoustic properties in words such as /atbros^j/ changed considerably from pre-training to post-training (i.e., F2 increased 298 Hz). While one of the trained palatalized consonants (viz., /brat^j/ showed at least a moderate increase in identification accuracy),the other trained

Table 6. Changes in identification accuracy among the 12 unpalatalized and palatalized target words from pre-training to post-training

	/atbros/	/atbros^j/	/brat/	/brat^j/	/kon/	/kon^j/	/top/	/top^j/	/udar/	/udar^j/	/ugel/	/ugel^j/
Pre-training	43	45	61	57	87	65	89	39	66	47	76	47
Post-training	43	50	63	71	89	87	85	65	64	51	81	62
Difference	0	5	2	14	2	22	−4	26	−2	4	5	15

palatalized word (/atbros^j/) showed very little improvement in identification accuracy from pre- to post-training. This was also the case with both of the trained unpalatalized targets (i.e., /atbros/ and /brat/).

To examine whether there was any relationship between the listeners' identification accuracy and the extent to which the 10 Russian learners distinguished between palatalized versus unpalatalized consonants, two correlation coefficients were calculated. The first was for listener identification accuracy for each of the L2 speakers versus each speaker's "difference value" between F2 frequency for vowels preceding palatalized and unpalatalized consonants at pre-training ($r = 0.91$, $p < .001$). The other comparison was for these same two measures at post-training ($r = 0.77$, $p < .01$). Thus, even though the various words the listeners heard were in random order across speakers, words and conditions, such that the listeners weren't able to directly compare minimal pairs of palatalized versus unpalatalized productions by a given speaker, their identification accuracy was highly correlated with the magnitude of the distinction individual learners made between such pairs of consonants.

6. Discussion

Hacking et al. (2016) confirmed that for native speakers of Russian, two of the most salient acoustic properties related to the production of word-final palatalized versus unpalatalized consonants were F2 frequency of the preceding vowel and final consonant "release noise." They found that experienced L2 Russian learners showed little, if any, distinction between palatalized and unpalatalized consonant targets in terms of these measures. The aim of the present study was to examine whether a training protocol, which included EPG feedback as its primary component, would lead to improvements in L2 speakers' abilities to produce a contrast between word-final palatalized versus unpalatalized consonant targets. To this end, the primary acoustic measures made in the present study concerned F2 frequency in vowels preceding palatalized versus unpalatalized consonant targets and post-consonant release noise.

Seven of the 10 L2 Russian learners showed statistically significant increases in F2 frequency of vowels preceding palatalized consonant targets when comparing their pre-training to post-training productions – ranging from 127 to 339 Hz. The other three L2 subjects also showed increases in the expected direction, but their increases were not statistically significant – ranging from 34 to 121 Hz. Because the training procedures were discontinued after the training period (i.e., at the conclusion of the semester), it is not known whether those three individuals (or any of the other students) might have achieved additional gains with more training. As noted earlier, the learners already evidenced a modest, but statistically significant difference between the F2 of vowels preceding palatalized and unpalatalized consonants at the pre-training evaluation. However, a comparison of their average F2 values at pre-training with native speaker values taken from Hacking et al. (2016) indicated that the learners' average pre-training F2 value preceding palatalized consonants was closer to the native speakers' F2 values for vowels preceding unpalatalized consonants. In addition, despite the fact that the learners showed a significant difference for this measure at pre-training, the contrast between F2 of vowels preceding palatalized versus unpalatalized consonants approximately doubled, for the group as a whole, from pre-training to mid-training to post-training.

Acoustic properties of consonant release noise have also been shown to be contrastive for native Russian speakers' palatalized versus unpalatalized consonants (Bolanos, 2013; Hacking et al., 2016; Ordin, 2010). The learners' pre- and post-training consonant release noise were compared to the findings for the native speaker data from the earlier study. The learners' release noise values were within the range of the values for the native speakers, and most of the learners' values did not change significantly with training. Furthermore, the pre-training data showed that even prior to training, the learners evidenced a statistically significant difference in release noise between the two types of consonants: palatalized versus unpalatalized. However, while the learners' values fell within the range produced by the native speakers and the difference between the palatalized and unpalatalized values was statistically significant, the learners' palatalized and unpalatalized production values were not as distinct, i.e., the distance between the two was not as great, as those of the native speakers. This suggests that what seems to be important is not just the absolute values of the respective consonant types but the magnitude of "difference" between the values of release noise for palatalized versus unpalatalized consonants.

Despite the statistically significant increases in F2 frequency of vowels preceding palatalized consonants from pre- to mid- and post-training, the native Russian listeners, as a group, showed only relatively small improvements in identification accuracy. This raises the question of how readily distinguishable are native productions of these same contrasts to native speaker listeners? While we did not include native productions in this listening task, prior research has shown that native speakers readily perceive the difference between unpalatalized and palatalized Russian consonants. In a gating experiment, Kavitskaya (2006: 589) showed that "cues for palatalization are at least as perceptually salient for speakers of Russian as cues for voicing and place of articulation." Babel and Johnson (2007), comparing the perception of palatalization contrasts for native Russian and native English speaking listeners, found that Russian native speakers were more sensitive to degrees of palatalization than were English native speakers.

It is possible that if the listening task had involved more direct comparisons (e.g., hearing two productions and judging whether they were the same or different), the native listeners may have been able to more readily distinguish contrasts made by the L2 speakers. The L1 listeners' relatively small increase in identification accuracy associated with hearing the various pre- and post-training productions in a random format seems to suggest that whatever improvements the L2 learners did make were often not substantial enough to exceed the absolute acoustic threshold required for productions to be identifiable. It is worth noting that these findings are consistent with earlier research on child disordered speech that documents sub-phonemic acoustic changes that do not result in perceptual discriminability by listeners (Gibbon, 1990; Gierut & Dinnsen, 1986).

Research on L2 phonological training often uses generalizability as a measure of efficacy; i.e., the degree to which learners are able to generalize learning to other contexts is an indicator of how robust a training protocol is (Hardison, 2004). For example, Hardy (1993) used the ability of an L1 Spanish subject to generalize L2 English phoneme learning from trained to untrained words as evidence of what she termed "ease of learning". More directly relevant to our training context is data on the subject's phonological restructuring. She showed that training resulted in progress toward achieving a phonological split between /d/ and /ð/. Because the learners in the present study received training on two pairs of sounds: /s/-/s^{j}/ and /t/-/t^{j}/ but were recorded producing words containing four additional pairs, viz. /l/-/l^{j}/, /n/-/n^{j}/, /r/-/r^{j}/, /p/-/p^{j}/, the data allow us to consider the issue of generalizability. Specifically, were learners able to generalize properties associated with the phonological contrast between unpalatalized and palatalized consonants from trained to untrained segments? As was observed, there was a statistically significant improvement in the F2 of vowels preceding the trained and the untrained palatalized targets from pre- to post-training. Among the six word pairs, there was only one pair (/udar/-/udar^{j}/) for which there was not a statistically significant improvement in F2 values. This is perhaps not surprising given that the palatalized trill involves "conflicting physical constraints on the tongue dorsum" (Kavitskaya et al., 2009), and /r^{j}/ is one of the last sounds to develop in the native speaking child's phonological inventory (Timm, 1977). However, in the aggregate, learners were able to generalize from the two trained consonant pairs to three of the four untrained consonant pairs.

A further contribution of this study pertains to training design. Maas et al. (2008) noted that there is no clear evidence in support of how many training sessions, of what duration, or over what length of time is most efficacious. Our protocol had students complete two two-week training periods, each consisting of eight 15-minute sessions. However, as reported above, on average, students showed significant increases in F2 preceding palatalized consonants between pre-training and mid-training, with limited additional changes between mid-training and post-training. This suggests that an intervention period of as little as eight 15-minute sessions over a period of two weeks can be effective in training at least certain contrasts. A further question is the role played by different types of feedback. Participants received knowledge of performance (KP) feedback in the form of some explicit feedback about tongue placement for the target consonants during the orientation sessions. They then monitored their progress against the representative EPG visual targets by tracking how dots lit up depending on tongue placement. Future research could further exploit KP feedback with more instruction on tongue physiology and mechanics as well as by incorporating dynamic EPG targets that show coarticulation, for example the movement of vowel into consonant.

Another goal of this study was to experiment with using EPG as part of a course and under a protocol that made students responsible for their own training. This is in contrast to clinical uses of EPG documented in the literature where training sessions are monitored by a clinician/researcher (McAuliffe & Cornwell, 2008). If we are to contemplate the broader use of EPG as a tool in L2 pronunciation training, we need to know whether students can use it independently and effectively. The results of this study are positive in this regard. After a general group orientation session and one short one-on-one session with the lead researcher, the rest of the training sessions were completed independently, suggesting that EPG is a viable training procedure for adult L2 learners. Unfortunately, EPG is currently a reasonably expensive technique. In addition to the initial outlay for software and datalinks that connect the pseudopalate and the computer (ca. $3000.00), each student's pseudopalate costs approximately $150, which could make it financially prohibitive. Nevertheless, in individual instances and/or as the technology becomes less expensive, it is encouraging to note that EPG may well benefit individuals attempting to improve their pronunciation of this, and potentially other, important phonological contrasts in Russian and other languages.

In addition to the financial cost, it is also important to note other possible limitations of the study. While it is the case that students showed improvement in their production of the contrast, we cannot say with certainty that this was entirely due to the EPG training. There was a brief listening component at the beginning of each training session, which could have contributed to at least a certain amount of the improvement that was observed. Furthermore, the study was conducted during a semester-long course on Russian phonetics and phonology, so students received instruction about Russian articulation, and the resulting increased awareness of the importance of this contrast could also have been a contributing factor to improvements that occurred. Finally, as already noted, because the study was conducted as part of a course, there was no formal control group, which makes it difficult to know whether an untrained group might also potentially have shown at least some improvement in the ability to produce the contrast simply as a function of factors other than the explicit training that these 10 learners were provided. However, it is important to note in this regard that subjects in two previous studies (Hacking, 2011; Hacking et al., 2016) were similar to the students in this study, in that they had completed four to six semesters of general Russian language training, including an advanced Russian phonetics and phonology course (but with no EPG and focused listening training). The (quasi-control) students in those studies showed little or no ability to produce a contrast between palatalized and unpalatalized consonants by the end of the course. Therefore, utilizing only explicit instruction about palatalization was not sufficient to improve production, whereas EPG training appears to be an effective intervention.

This study showed that a group of learners who completed EPG training on the Russian palatalization contrast exemplified in two consonant pairs – /t/-/t^j/ and /s/-/s^j/ – made statistically significant improvements in the acoustic measures associated with this important Russian phonological contrast. They were also able to generalize these improvements to three of four other consonant pairs on which they did not train, a fact which speaks to the robustness of EPG training for this particular phonological feature. However, these improvements in production did not result in substantial gains in identification of the target segments by native Russian listeners. While this is a somewhat discouraging finding, the fact that there was some improvement in the listener data from pre- to post-training is important. The combination of an increase in perceptibility, however small, and the evidence of acoustic changes in the learners' productions is consistent with phonological restructuring in process. Further research is needed to determine how EPG training might be adjusted to produce further gains.

References

Adams, S. G., & Page, A. D. (2000). Effects of selected practice and feedback variables on speech motor learning. *Journal of Medical Speech-Language Pathology*, 8, 215–220.

Antonova, D. N. (1988). *Fonetika i intonacija*. Moscow: Russkij jazyk.

Avanesov, R. I. (1972). *Russkoe literaturnoe proiznošenie*. Moscow: Prosveščenie.

Babel, M., & Johnson, K. (2007). Cross-linguistic differences in the perception of palatalization. Proceedings of the 16th International Congress of the Phonetic Sciences.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer* [Computer program, version 5.3.23]. Retrieved from: <http://www.praat.org/>

Bolanos, L. (2013). Perception and production in non-native speech: Russian palatalization. Proceedings of *Meetings on Acoustics. Acoustical Society of America*, 19, 1–7.

Bryzgunova, E. A. (1963). *Praktičeskaja fonetika i intonacija russkogo jazyka*. Moscow: Izdatel'stvo moskovskogo universiteta.

Cleland, J., Timmins, C., Wood, S. E., Hardcastle, W. J., & Wishart, J. G. (2009). Electropalatographic therapy for children and young people with Down's syndrome. *Clinical Linguistics & Phonetics*, 23, 926–939. doi: 10.3109/02699200903061776

Demenko, G., Wagner, A., & Cylwik, N. (2010). The use of speech technology in foreign language pronunciation training. *Archives of Acoustics*, 35, 309–329.

Diehm, E. E. (1998). *Gestures and linguistic function in learning Russian: Production and perception studies of Russian palatalized consonants*. Doctoral dissertation, Ohio State University. Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 9833968).

Engwall, O. (2012). Analysis of and feedback on phonetic features in pronunciation training with a virtual teacher. *Computer Assisted Language Learning*, 25, 37–64. doi: 10.1080/09588221.2011.582845

Fant, G. (1970). *Acoustic theory of speech production*. The Hague: Mouton De Gruyter.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116–124. doi: 10.3758/BF03195503

Fujiwara, Y. (2007). Electropalatography home training using a portable training unit for Japanese children with cleft palate. *International Journal of Speech-Language Pathology*, 9, 65–72. doi: 10.1080/14417040601120904

Gibbon, F. (1990). Lingual activity in two speech-disordered children's attempts to produce velar and alveolar stop consonants: Evidence from electropalatographic (EPG) data. *International Journal of Language & Communication Disorders*, 25, 329–340. doi: 10.3109/13682829009011981

Gibbon, F. E., & Wood, S. E. (2003). Using electropalatography (EPG) to diagnose and treat articulation disorders associated with mild cerebral palsy: a case study. *Clinical Linguistics & Phonetics*, 17, 365–374. doi: 10.1080/0269920031000079976

Gibbon, F. E., Hardcastle, W. J., & Suzuki, H. (1991). An electropalatographic study of the /r/, /1/ distinction for Japanese learners of English. *Computer Assisted Language Learning*, 4, 153–171. doi: 10.1080/0958822910040304

Gierut, J. A., & Dinnsen, D. A. (1986). On word-initial voicing: Converging sources of evidence in phonologically disordered speech. *Language and Speech*, 29, 97–114.

Hacking, J. (2011). The production of palatalized and unpalatalized consonants in Russian by American learners. In Magdalena Wrembel, Małgorzata Kul, & Katarzyna Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds* (pp. 93–101). Frankfurt: Peter Lang.

Hacking, J., Smith, B., Nissen, S., & Allen, H. (2016). Russian palatalized and unpalatalized coda consonants: An electropalatographic and acoustic analysis of native speaker and L2 learner productions. *Journal of Phonetics*, 54, 98–108. doi: 10.1016/j.wocn.2015.09.007

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology*, 8(1), 34–52.

Hardy, J. E. (1993). Phonological learning and retention in second language acquisition. In F. R. Eckman (Ed.), *Confluence: Linguistics, L2 acquisition and speech pathology* (pp. 235–248). Amsterdam: John Benjamins. doi: 10.1075/lald.4.17har

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America*, 138, 817–832. doi: 10.1121/1.4926561

Kavitskaya, D. (2006). Perceptual salience and palatalization in Russian. *Laboratory phonology* 8, 589–610.

Kavitskaya, D., Iskarous, K., Noiray, A., & Proctor, M. (2009). Trills and palatalization: Consequences for sound change. *Proceedings of the Formal Approaches to Slavic Linguistics*, 17, 97–110.

Kochetov, A. (2006). Testing licensing by cue. *Phonetica*, 63, 113–148. doi: 10.1159/000095305

Lohmander, A., Henriksson, C., & Havstam, C. (2010). Electropalatography in home training of retracted articulation in a Swedish child with cleft palate: Effect on articulation pattern and speech. *International journal of speech-language pathology*, 12, 483–496. doi: 10.3109/17549501003782397

Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. *American Journal of Speech-Language Pathology*, 17, 277–298. doi: 10.1044/1058-0360(2008/025)

Machovikov, A., Stolyarov, K., Chernov, M., Sinclair, I., & Machovikova, I. (2002). Computer-based training system for Russian word pronunciation. *Computer Assisted Language Learning*, 15, 201–214. doi: 10.1076/call.15.2.201.8192

Massaro, D. W., Liu, Y., Chen, T. H., & Perfetti, C. (2006). A multilingual embodied conversational agent for tutoring speech and language learning. Proceedings of *Interspeech* 2006 (pp. 825–828).

McAuliffe, M. J., & Cornwell, P. L. (2008). Intervention for lateral/s/using electropalatography (EPG) biofeedback and an intensive motor learning approach: a case report. *International Journal of Language & Communication Disorders*, 43, 219–229. doi: 10.1080/13682820701344078

Nordberg, A., Carlsson, G., & Lohmander, A. (2011). Electropalatography in the description and treatment of speech disorders in five children with cerebral palsy. *Clinical Linguistics and Phonetics*, 25, 831–852. doi: 10.3109/02699206.2011.573122

Ordin, M. (2010). Palatalization and temporal organization of CVC clusters in Russian. *Russian Linguistics*, 34, 57–65. doi: 10.1007/s11185-009-9049-4

Patten, I., & Edmonds, L. A. (2015). Effect of training Japanese L1 speakers in the production of American English /r/ using spectrographic visual feedback. *Computer Assisted Language Learning*, 28, 241–259. doi: 10.1080/09588221.2013.839570

Purcell, E.T. (1979). Formant frequency patterns in Russian VCV utterances. *Journal of the Acoustical Society of America*, 66, 1691–1702. doi: 10.1121/1.383641

Ruellot, V. (2011). Computer-assisted pronunciation learning of French /u/ and /y/ at the intermediate level. In Proceedings of the *2nd Pronunciation in Second Language Learning and Teaching Conference* (pp. 199–213).

Schmidt, A. M. (2012). Effects of EPG treatment for English consonant contrasts on L2 perception and production. *Clinical Linguistics & Phonetics*, 26, 909–925. doi: 10.3109/02699206.2012.718036

Schmidt, A. M., & Beamer, J. (1998). Electropalatography treatment for training Thai speakers of English. *Clinical Linguistics & Phonetics*, 12, 389–403. doi: 10.1080/02699209808985233

Smith, B. L., Hillenbrand, J., & Ingrisano, D. (1986). A comparison of temporal measures of speech using spectrograms and digital oscillograms. *Journal of Speech, Language, and Hearing Research*, 29, 270–274. doi: 10.1044/jshr.2902.270

Timm, L. A. (1977). A child's acquisition of Russian phonology. *Journal of Child Language*, 4, 329–339.

Authors' addresses

Jane F. Hacking
Languages and Literatures
255 S. Central Campus Dr.
LNCO 1930
Salt Lake City, UT 84112
USA

Bruce L. Smith
Communication Sciences and Disorders
390 S. 1530 E.
Salt Lake City, UT 84112
USA

Eric M. Johnson
2416 South 800 East
Salt Lake City, UT 84106
USA