City University of New York (CUNY)

CUNY Academic Works

Publications and Research

Kingsborough Community College

2012

Palatalization in Romanian — Acoustic properties and perception

Laura Spinu

CUNY Kingsborough Community College

Irene Vogel

University of Delaware

H. Timothy Bunnell

Nemours Biomedical Research

[How does access to this work benefit you? Let us know!](#)

More information about this work at: https://academicworks.cuny.edu/kb_pubs/162

Discover additional works at: <https://academicworks.cuny.edu>

This work is made publicly available by the City University of New York (CUNY).

Contact: AcademicWorks@cuny.edu

Palatalization in Romanian - Acoustic Properties and Perception

Laura Spinu^a Irene Vogel^b H. Timothy Bunnell^c

^aDepartment of Classics, Modern Languages and Linguistics

Concordia University

1455 de Maisonneuve Blvd. West, H-1013-4

Montreal, QC, Canada H3G 1M8

Tel. 514-848-2424 ext. 5159

Fax 514-848-8679

lspinu@alcor.concordia.ca

^bUniversity of Delaware

^cUniversity of Delaware and Nemours Biomedical Research

Abstract

This paper presents the results of an acoustic study of fricatives from four places of articulation produced by thirty-one native speakers of Romanian, as well as those of a perceptual study using the stimuli from the acoustic experiment, allowing for a direct comparison between acoustic properties and perception. It was found that there are greater acoustic differences between plain and palatalized labials and dorsals as compared to coronals. The acoustic results were paralleled by the perceptual findings. This pattern departs from cross-linguistic generalizations made with respect to the properties of secondary palatalization. A likely source of the differences is the fact that previous studies of secondary palatalization typically involved stops which tend to exhibit various enhancement phenomena at the coronal place of articulation. Since the enhancement generally involves additional frication, this is not a useful strategy for fricatives at the coronal, or any other place of articulation. These findings form the basis of a discussion highlighting the differences between enhanced and non-enhanced secondary palatalization.

Key words: secondary palatalization, fricatives, Romanian, perception, cepstral analysis, place of articulation, enhancement

1 Introduction

While palatalization phenomena are common in Romance languages, Romanian stands out in exhibiting an additional phenomenon of *secondary* palatalization. More specifically, while full palatalization¹, also known as coronalization (Hume 1994, Bateman 2007), is frequently encountered in other Romance languages, secondary palatalization² is a phenomenon generally considered characteristic of Celtic and Slavic languages, among others, where the distinction between plain consonants and those with secondary palatalization is contrastive. Secondary palatalization has been described in detail in Russian and other Slavic languages (Fant 1970, Kenstowicz 1972, Comrie 1981, Ladefoged 1993, Kochetov 2002), as well as in Irish (Ní Chiosáin 1994). Cross-linguistic studies have revealed several generalizations with respect to the perceptual salience of secondary palatalization at different places of articulation, specifically higher perceptual salience of secondary palatalization was reported for coronals as compared to labials (Kochetov 2002, Kavitskaya 2006). By contrast, native speakers of Romanian were found to be more sensitive to the plain-palatalized contrast in labials (Spinu 2007, 2009). As the manner of articulation of the consonants examined differed across these studies (stops in Kochetov and Kavitskaya’s studies, fricatives in Spinu’s), this suggests that the perceptual properties of secondary palatalization are not uniform across different manners. The goal of this paper is to explore the reasons underlying these differences in more detail.

To do so, we present the findings of a production experiment in which we elicited plain and palatalized consonants in Romanian and examined their acoustic properties. We furthermore conducted a perceptual experiment employing the stimuli elicited in the acoustic study, in order to be able to make a direct comparison between the acoustic and perceptual properties of these segments.

In the following sections, we first describe the phenomenon of secondary palatalization in Romanian. We summarize the findings of the previous perceptual studies of secondary palatalization in Romanian, and compare these findings with reported cross-linguistic generalizations (Section 2). We then present our production experiment and its results, which enable us to make predictions as to the perception of this contrast at various places of articulation (Section 3). Next, our perception experiment is presented in Section 4. In Section 5, we discuss the overall results in relation to the more general question of secondary palatalization across languages, and the potential contribution of phonetic enhancement to the reported properties of the plain-palatalized

¹ This process refers to the change in a (usually velar) consonant’s primary place of articulation to coronal.

² Secondary palatalization refers to the presence of a secondary palatal feature accompanying the primary place of articulation of a consonant.

contrast. Finally, we summarize our findings in the conclusion section.

2 Secondary Palatalization in Romanian

2.1 *The phenomenon of secondary palatalization*

While secondary palatalization in a language such as Russian is a reflection of an underlying contrast between plain and palatalized consonants (Bhat 1978, Padgett 2001, Bateman 2007), in Romanian the surface contrast arises as the result of a rule operating in certain phonological and morphological contexts, and thus is not present underlyingly³ (Chitoran 2002, Iscrulescu 2003). Thus, in Russian we find minimal pairs such as *[p]at* 'stalemate' / *[pʲ]at* 'heel-gen.pl.', but in Romanian we most often find surface palatalized consonants associated with the presence of one of two homophonous inflectional suffixes consisting of /-i/: the plural for nouns and adjectives (1a), and the 2nd person singular, present indicative of verbs (1b). The /-i/ itself is not pronounced, but is manifested as palatalization of the previous consonant - characterized by a secondary palatal gesture of the tongue body accompanying the primary articulation of the respective consonant.

- (1) a. [domn] 'gentleman' vs. [domnʲ] 'gentlemen' UR: /domn+i/
 b. [sar] 'I jump' vs. [sarʲ] 'you jump' UR: /sar+i/

Whereas the occurrence of secondary palatalization in Romanian most commonly coincides with the morphological contexts described above, there are also monomorphemic items that exhibit final secondary palatalization in the absence of the usual morphemes, as shown in (2). In these cases, the lexical item is taken to end in an underlying /i/.

- (2) a. [ungʲ] 'angle' (pl. [ungjurʲ]) b. [baremʲ] 'at least' (adv.)
 c. [puftʲ] 'kid' d. [azʲ] 'today'

Despite the fact that secondary palatalization occurs with the vast majority of items ending in /Ci/, it should be noted that there are certain cases in which

³ A certain degree of morphological conditioning, however, is present in Russian as well, as word-final palatalized consonants can also be markers of certain morphological classes, such as gender or declension.

it fails to apply and instead the full [i] vowel appears. These cases, however, are quite limited, and are morphologically and phonologically predictable, for example in monosyllabic words ending in /i/, (e.g. *zi* [zi] 'day', *gri* [gri] 'gray') (cf. Vogel and Spinu 2009).

Given that secondary palatalization in Romanian depends on the presence of /-i/ after a consonant at the end of a word, we observe this phenomenon only in (word-final) codas.

2.2 Perception of secondary palatalization

Based on a survey of twenty-two languages from the Slavic, Celtic, and Uralic families, Kochetov (Kochetov 2002) presents a set of cross-linguistic asymmetries relating to secondary palatalization. He concludes that palatalized labials are more marked than palatalized coronals on the basis of several distributional patterns (e.g. palatalized labials are less frequent than palatalized coronals throughout the world's languages, and the plain-palatalized contrast with labials is the first target for neutralization in disfavored environments such as the coda position). In a similar, but more large-scale, survey, Bateman (2007) also uncovered a set of implicational relations whereby the secondary palatalization contrast with labials can only be found in languages that also have palatalization with coronals, or dorsals, or both.

The distributional patterns are supported by Kochetov with a series of perceptual experiments. Specifically, he found that Russian and Japanese listeners were more successful in identifying plain and palatalized coronals, as opposed to labials, when these were presented in nonce words produced by a Russian male speaker (Kochetov 2002). This place of articulation asymmetry was also observed in a gating experiment with 10 native speakers of Russian (Kavitskaya 2006), as palatalization cues were missed in the case of palatalized labials significantly more often than with palatalized coronals.

The pattern observed by Kochetov failed to be replicated in three experiments with Romanian listeners which used real words of Romanian produced by a male speaker of this language (Spinu 2007, 2009). In Spinu (2007) the first experiment (n=20) was an identification task in which the items appeared out of context in carrier sentences, as illustrated in (3), while the second experiment (n=12) involved a forced choice grammaticality judgment task in which the targets were presented in sentences with matched and mismatched morphological cues, as illustrated in (4). In (4a), the plural cue (*patru* 'four') is matched with the presence of palatalization (*ciorapⁱ* 'socks') which, as discussed in Section 2.1, represents the plural marker. By contrast, in (4b) the same plural cue is mismatched with the noun (*ciorap* 'sock') as the absence of palatalization indicates singular status.

- (3) Am să aleg cuvântul [lup]/[lup^j] mâine.
 'I will choose the word 'wolf/wolves' tomorrow.'

(4) a. **Matched**

După ce am spălat rufe nu am mai găsit patru ciora[p^j].
 'After doing the laundry, I couldn't find four socks.'

b. **Mismatched**

După ce am spălat rufe nu am mai găsit *patru ciora[p].
 'After doing the laundry, I couldn't find *four sock.'

The listeners showed similar response patterns in both experiments: they were slightly more successful with the plain-palatalized contrast in [p] than in [ts] (the difference did not reach statistical significance), and both these results were significantly better than those with /f/.

While the difference between labials and [+anterior] coronals did not reach statistical significance in these two experiments, a third experiment with twenty-two subjects (Spinu 2009) does show such a difference. That is, the Romanian speakers exhibited higher sensitivity to the plain-palatalized contrast with labials versus coronals. It should be noted that a generalization as to place of articulation could not be drawn based on the first two experiments, given that manner of articulation also differed across the three consonants tested ([p], [ts], and [f]). This problem was avoided in the third experiment which considered only fricatives⁴. The results of this study revealed that subjects' sensitivity to palatalization with labial /v/ was significantly higher than with

⁴ The reason a stop series was not investigated is because in Romanian a root-final [+anterior] coronal stop involves further changes (e.g. assibilation) in the presence of the plural morpheme, aside from acquiring a secondary palatal feature:

- a. po[t] 'I can' vs. po[ts^j] 'you can'
 b. plo[d] 'child' vs. plo[z^j] 'children'

As for velar stops, secondary palatalization is generally accompanied by coronalization.

- c. ra[k] 'crawfish-sg.' vs. ra[tʃ^j] 'crawfish-pl.'

Thus, the comparison is in effect between two segments that differ in more than just the presence/absence of the secondary palatalization gesture, specifically in primary place of articulation, release burst, as well as secondary palatalization. As a result, the contrast becomes highly salient, but the salience is not due to the palatalization per se, but to additional differences between the plain and palatalized form of a consonant. This would have not fitted the goal of examining the properties of secondarily palatalized consonants in the absence of such differences, so as to be able to draw generalizations regarding the interaction of secondary palatalization with different places of articulation (all other things being equal, to the extent possible).

coronal /z/, $t(21)=2.42$, $p<.05$.

The fact that the results for Romanian were consistently different from the cross-linguistic generalizations, as well as from previous perceptual findings, raises interesting questions with regard to the properties of secondary palatalization in general. We thus conducted a production study, discussed in the next section, to determine the acoustic properties of Romanian plain and palatalized consonants at different places of articulation. Our study also expands on the previous research, by examining more places of articulation.

3 Romanian production experiment

3.1 *Experimental design*

3.1.1 *Subjects and procedure*

Thirty-one native speakers of standard Romanian participated in the experiment. The subjects were 10 males and 21 females, ranging in age from 19 to 30, with an average age of 21.7 years. They were all tested individually in a quiet room in Bucharest, Romania, and were compensated for their time. The stimuli were presented and recorded using the InvTool program (InvTool 2007). Specifically, each sentence was displayed on a computer screen (in orthography) and the subjects were instructed to read the sentence in a natural-sounding way. The program prompted for repetitions of the same sentence if it was either too loud or not sufficiently loud or if the fluctuations in pitch were too large, but no manipulation of the sound files took place. If no problems were encountered, the utterance was saved and the next sentence appeared on the screen. Before beginning the actual experiment, the subjects had a practice session with 20 items to familiarize themselves with the procedure.

3.1.2 *Stimuli*

The stimuli consisted of four pairs of words ending in a plain and a palatalized version of the following five fricatives: /f, v, z, ʃ, h/. The choice of fricatives has to do, on the one hand, with limitations of the Romanian consonant inventory (fricatives outnumber stops in terms of places of articulation available), and on the other hand, with the avoidance of pairs of segments that differ in more than just the presence/absence of a secondary palatal gesture, by displaying assibilation, change of primary place of articulation, or other possible enhancement effects, as is the case with coronal or velar stops (see Footnote 4). Thus, four places of articulation are considered (the place features listed below are based on Clements and Hume 1995):

[Labial]: /f/ and /v/ have traditionally been described as labial (labio-dental) (Chitoran 2002)

[Coronal, +anterior]: /z/ in Romanian is classified as dental (Chitoran 2002)

[Coronal, -anterior]: /ʃ/ is referred to as post-alveolar or palatal (Chitoran 2002)

[Dorsal]: /h/ - descriptions of the phonetic value of orthographic 'h' have been vague. Most often, it has been referred to as a glottal fricative (Chitoran 2002), but also as a glide (Ruhlen 1973), with the mention that it can be realized as a velar fricative, particularly before liquids or in word-final position (Mallinson 1986). As only word-final segments are included in the experiment, this segment will be referred to as dorsal.

In the remainder of the paper, when referring to the specific consonants under investigation, we will be using the phonemic transcription, that is /.../, so as to avoid committing to a specific representation as far as their surface forms are concerned. This is particularly relevant regarding the dorsal segment, for which we will be using the symbol /h/. According to the International Phonetic Alphabet (IPA), this symbol corresponds to a glottal fricative. We will thus be disregarding the possible allophony with the velar fricative [x], which to date has not been confirmed by any acoustic studies. For the time being, it is not the precise articulation that is of interest, but rather the addition of a new place to the categories already explored, as no place further back than [+anterior] coronal was included in the previous acoustic and perceptual studies (Kochetov 2002, Kavitskaya 2006).

The Romanian inventory includes two additional fricatives, specifically [s] and [ʃ]. The former was not included in this study because, together with [d], it is one of the only two consonants in the Romanian inventory that never appear in palatalized form on the surface (as shown in Footnote 4, [d] alternates with [z] in the plural. Similarly, [s] alternates with [ʃ]). As for [ʃ], its palatalized form is rarely encountered in Romanian, and thus not enough target words of the required shape could be found.

For all fricatives tested in this experiment, the difference between the plain and palatalized form is assumed to reside in the absence versus presence of the additional palatal gesture, with no further changes in place or manner of articulation (as is the case with stops). The choice of fricatives also permits us to address the question of whether the asymmetries associated with the palatalization contrast with stops might hold for any manner of articulation (Kochetov 2002, p. 51).

Thus, the stimuli consisted of pairs of words that differed only in whether their final consonant was plain or secondarily palatalized. They were presented in a context-neutral carrier sentence, as shown in (5). The targets are shown in phonetic transcription inside square brackets (the full set of target words can be found in the appendix). All of the target words were disyllabic, with the stress falling on the final syllable.

(5) Am să aleg cuvântul [pantof/pantof^ɨ] când voi fi gata.

‘I will choose the word [shoe/shoes] when I am ready.’

It should be noted that the material following the target (*când*, [kin(d)]) had previously been determined in a pilot study not to induce anticipatory co-articulation effects on the target consonant, in contrast with a following labial or coronal stop. A following vowel would not have been feasible, due to the phenomenon of coda resyllabification typically encountered in Romanian and other Romance languages, whereby a word-final consonant becomes an onset to a following vowel (in the case of word-final secondary palatalization, the palatal element might be resyllabified as a glide onset).

In addition to the target sentences, the subjects produced twice as many fillers. The fillers were also paired, but showed inflections other than singular or plural, so as to distract the subjects from the target pairs. For example, one of the filler pairs was *oceanul* ‘the ocean’ and *oceanic* ‘oceanic’. The set of 120 items (40 targets + 80 fillers) was presented three times to each subject, with the items automatically randomized for each block. Thus, a full set of recordings contained 120 target items per subject: 5 consonants x 4 words per consonant x 2 forms per item (plain and palatalized) x 3 repetitions. All of the subjects except one produced the full set of three repetitions; the remaining subject only produced two repetitions.

3.2 Data analysis

Six items were rejected due to disfluencies, leaving 3,674 items for acoustic analysis. Each segment was acoustically analyzed to obtain the duration and average spectral properties expressed as the first six coefficients of the Bark cepstrum, which describes the amplitude and shape of the speech spectrum in terms of a set of compact orthogonal components (Bunnell et al. 2004).

The choice of cepstral coefficients was motivated primarily by the fact that this study is concerned with the *contrast* between plain and palatalized members of a pair, and not each member of a pair separately. As we will see, the cepstral method is particularly suitable for this situation. Moreover, cepstral coefficients present advantages over comparable methods such as spectral moments because of their relationship with the speech spectrum. Another advantage of using this method has to do with the possibility of automatizing the extraction of the acoustic values. Thus, very little manual work is required compared to more traditional measures, which makes it suitable for the analysis of large corpora such as the one obtained for the experiment presented here. Finally, cepstral methods have been successfully applied in phonetic research in the past. Each of these reasons is elaborated on in the following paragraphs.

With respect to the analysis of a contrast between two sounds, cepstral coeffi-

coefficients offer not only a way to quantify the acoustic differences between the two, but also the possibility to compare them to those of other pairs. Each of the six coefficients that we used is characterized by one value. By comparing the values for the plain versus the palatalized member of a pair, we can assess the acoustic distance for that particular contrast (how similar/different the two sounds are with respect to each other), regardless of the cues that may play a role in perception. For instance, spectral peak location has been shown to be an important acoustic cue for sibilant fricatives such as [s] or [ʃ], whereas other cues were found useful in describing non-sibilant fricatives (e.g. flat spectral energy distribution in the case of labial fricatives, or the presence of several formant-like peaks for dorsal fricatives). As cepstral coefficients capture information about the overall spectrum shape, which cues are most important to the description of any particular sound (or contrast) does not have any bearing on the results. Once it is established which of the coefficients has proven most useful in the identification of a given sound it becomes possible, however, to explore the ways in which it correlates with more traditional cues.

Just like cepstral coefficients, the spectral moments analysis also captures features of overall spectrum shape in a compact number of coefficients. However, given the computation of the cepstral coefficients and their relationship with the speech spectrum, they are likely to provide more information in the future as far as articulatory details are concerned, as opposed to comparable analyses like spectral moments which are based on purely statistical concepts such as mean, variance, skewness and kurtosis. Figure 1 shows the six cepstral feature vectors used in this study. The corresponding cepstral coefficients are the sum of the product of the cepstral feature vectors and the speech spectrum.

Thus, c_0 is the sum of all the spectral values multiplied by 1.0, i.e., the total spectral energy. With respect to c_1 , the sum of the product of that vector with a speech spectrum will be large and positive when the spectrum has strong low frequency energy relative to the high frequency energy (e.g. a voiced segment). For a segment like /s/ that has little low frequency energy and strong high frequency energy, c_1 will have a negative value because the part of the vector with negative values will be multiplied against the strongest part of the spectrum. For c_2 , the coefficient will be most strongly positive for segments that have strong low and high frequency energy and very weak energy in the mid frequency regions. Alternatively, for a segment that has strong mid-frequency energy and weak low and high frequency energy, the c_2 coefficient will be large and negative. In sum, each coefficient will be large and positive when the speech spectrum looks just like the feature vector. The coefficient will be large and negative when the speech spectrum is the complement of the feature vector. As the computation of each coefficient is based on the energy distribution in certain frequency regions in the spectrum (e.g. low and high frequencies for c_1 , low, mid, and high frequencies for c_2), it is hoped that in the future, together with a clearer understanding of articulatory details, they may bring along new information concerning the relationship between articulation and specific frequency regions.

Fig. 1. Cepstral feature vectors 0 through 5. Note that these coefficients are drawn here on a linear frequency axis, but the frequency axis was Bark-scaled (see Footnote 5) for the analyses presented in this chapter.

Another reason for choosing the cepstral analysis has to do with the particularly large corpus analyzed here. An important advantage of using perceptually weighted (in this study, Bark-scaled⁵) cepstral coefficients is that their computation is straightforward and unambiguous. The more traditional alternative for acoustic analysis of fricatives would have been to find formant-like peaks in the spectrum and look for differences in formant frequencies associated with the plain/palatalized distinction, but this would not have been feasible

⁵ Bark scaling consists of passing the log magnitude spectrum through a large number of overlapping Bark-width bandpass filters, which has the effect of compressing the spectrum at higher frequencies and expanding it at lower frequencies in a way that corresponds to human auditory system processes.

for 3,674 utterances. In the case of a cepstral analysis, the only measurement that requires manual verification is the segmentation (which is the case for all other measurements discussed), but no other visual inspection is necessary, such as analysis of fricative spectra (to confirm spectral peak location), etc. The third, and crucial, reason for this choice is that previous studies have shown cepstral coefficients to be a successful, reliable, and completely reproducible method in phonetic research. Thus, the results of a large-scale study showed that better classification of voiceless stop release bursts (from a database of children’s speech containing recordings from 208 children) was obtained using a number of Bark cepstrum coefficients compared to the same number of spectral moments (Bunnell et al. 2004). Since the focus in the current study is on fricatives, which are similar to bursts in that they are both characterized by noise, it is important to choose a measure that was found to be reliable with this type of sounds. Moreover, a recent study (Ferragne and Pellegrino 2010) employing a novel methodology that computes distances between vowels in the Mel-Frequency Cepstral Coefficient (MFCC) space yielded a very promising estimate of the acoustic distance between accents (13 accents of the British Isles were investigated). The authors conclude that, "granted that distances in the MFCC space achieve good phonetic interpretability, [...] the argument that MFCCs cannot be wrong (while formants can) provides strong support for the use of MFCCs in phonetic studies, if only for practical reasons".

Returning to our experiment, the phonetic symbols were first automatically aligned by InvTool with the corresponding segments in the recordings. Both plain and palatalized consonants were treated as single segments to permit direct comparison, although the latter are represented in Romanian orthography as a sequence (C+i). Each file (i.e. recorded target sentence) was then examined to verify that the alignments were correct; any errors were corrected manually. Both the waveform and the wideband spectrogram of each token were used in verifying the segmentation. Fricative onset was defined as the point at which high-frequency energy first appeared on the spectrogram, and/or the point at which the number of zero crossings rapidly increased (Jongman et al. 2000). The end of each segment (plain or palatalized) was defined as the intensity minimum immediately preceding the silence during the closure of the following stop ([k] in all cases). The durations of the target consonants were determined, and then a Bark cepstrum analysis was performed in which six cepstral coefficients (DC and the first five cosine terms) were estimated separately for all frames of 10 ms of each segment. Average values were then computed for each segment.

Two types of statistical analyses were then conducted with the data. First, each acoustic measure was used as a dependent variable in a series of repeated measures, within-subjects analyses of variance (ANOVAs) with the independent factors consonant (Con) and palatalization status (Pal). Second, a stepwise Linear Discriminant Analysis (LDA), with leave-one-out cross-validation and palatalization (plain/palatalized) as the grouping variable was also performed

Fig. 2. Mean duration for plain and palatalized consonants. The symbol S stands for the post-alveolar fricative /ʃ/.

using all of the acoustic coefficients and duration to separate the plain and palatalized classes in a statistically optimal way.

3.3 Results

In the ANOVAs, the dependent variables were *duration* and the cepstral coefficients c_0 , c_1 , c_2 , c_3 , c_4 , c_5 ; the independent variables were Con (/f, v, z, ʃ, h/), and Pal (plain/palatalized). Significant main effects were found for Con and Pal on all the dependent variables; significant interactions between these two factors were also observed in all cases. That is, the average values for duration and the cepstral coefficients differed significantly on the basis of the consonant involved and whether or not it was palatalized. The interaction of these items also yielded significant differences.

The results for duration show that, while palatalized consonants were always longer than the plain ones, the difference was only significant ($p < .05$) in the case of /h/ and /v/. Figure 2 displays the mean duration values (in ms) for all the plain and palatalized segments. As can be seen, /v/ and /h/ were the consonants with the largest difference between the plain and palatalized form. The results for the different cepstral coefficients are very similar, in that larger differences can be noted between plain and palatalized labials, and especially dorsals, while smaller (if any) differences are found between the plain and palatalized coronals. This, as well as the ways in which the different types of consonants can be distinguished from each other (for instance, the voicing distinction between /f/ and /v/ is captured by differences in c_0 , c_1 , and c_2 , but not the other coefficients), can be seen in Figure 3.

Fig. 3. Mean cepstral coefficients for plain and palatalized consonants. The continuous lines correspond to the plain consonants and the dashed lines to the palatalized consonants. The error bars represent 95 % confidence intervals. The symbol S stands for the post-alveolar fricative /ʃ/.

With regard to the LDA, 78 % of the cross-validated grouped cases were classified correctly, but there were large differences depending on the specific consonant. The percentage of correct classifications for each segment, plain and palatalized, is shown in Table 1. As can be seen, a large discrepancy is present between the classification of plain [ʃ] (correctly classified 19.5% of the time) and its palatalized counterpart, (correctly classified 96.7% of the time).

	PLAIN	PALATALIZED
f	98.3	66.5
v	85	82
z	69	66.8
ʃ	19.5	96.7
h	99.4	97.5

Table 1

Correct LDA classification (%) for each segment, plain and palatalized.

This pattern, as we will see, was also paralleled in human perception (see Section 4). We return to this point in our discussion (Section 5.1).

It should be noted that these percentages include all of the elements examined, however, the strongest contributors to the LDA model were found to be c4 and c3. The contribution of each predictor variable to the LDA was assessed based on the standardized canonical discriminant function coefficients, which showed that the main three contributors were c4, c3, and c5 (this is further seen in Table 2, with c4 and c3 differing significantly between the plain and palatalized form of four of the five pairs, and c5 for three of the pairs). Subsequent discriminant analyses including only certain combinations of these predictor variables yielded slightly lower accurate classifications of plain and palatalized items, e.g. 75.9 % (for c4 and c3 only), 76.9% (for c4, c3, and c5), 76.8 % (for c4, c3, c5, and c1), and 77.2 % (for c4, c3, c5, c1, and duration). Regardless of the combination of predictors, the overall accuracy for classification of plain and palatalized consonants remained quite high, above 75% in all cases, and the combination of the two strongest predictors only lowered the overall accuracy by about 2% as compared to the LDA using all 7 independent variables as predictors.

Table 2 summarizes which acoustic measures varied significantly (at the .05 level) for the plain vs. palatalized form of each consonant, which can also be seen in Figures 2 and 3. As can be seen, plain vs. palatalized /h/ was the most successfully discriminated pair, with significant differences in all measures except c0; in fact, c0 is the only coefficient that never differed significantly between the plain and the palatalized form of a consonant. By contrast, plain vs. palatalized /ʃ/ could not be accurately discriminated, as shown by the fact that none of the acoustic measures was found to differ significantly between the plain and the palatalized form.

3.4 Summary and predictions

The results of the current experiment show that the primary place of articulation plays a crucial role in the acoustic properties of plain vs. palatalized

	SIGNIFICANT DIFFERENCES PLAIN VS. PALATALIZED
f	c1, c2, c3, c4, c5
v	duration, c1, c3, c4, c5
z	c3, c4
ʃ	-
h	duration, c1, c2, c3, c4, c5

Table 2

Significant differences in acoustic measures between the plain and palatalized form of each consonant.

consonants in Romanian, just as the previous perceptual results suggested. Specifically, greater differences between plain and palatalized segments appear in non-coronals (labials /v/ and /f/, and dorsal /h/) rather than in coronals (/z/ and /ʃ/).

While previous studies did not address the properties of secondary palatalization in dorsals, the present results reveal that in Romanian palatalization is most acoustically distinct at this place of articulation, with 6 of the 7 acoustic measures investigated differing significantly between the plain and palatalized form of /h/.

Based on the findings reported in this section, certain predictions can be made with respect to the perception of the consonants under investigation. It should be made clear that we do not expect human perception to replicate exactly the behavior of the LDA, and the reader must bear in mind that while the LDA operated strictly on the basis of durational and cepstral properties of the fricatives, additional cues are likely to play a role in human perception, for instance, vocalic transitions. Some very strong patterns, however, were identified, that we expect to find with human listeners as well. Specifically, given the lack of significant acoustic differences between plain and palatalized /ʃ/, (1) we expect listeners to do poorly in distinguishing this contrast, and furthermore, to display the same bias for hearing plain /ʃ/ as palatalized but not vice versa; (2) dorsal /h/ is likely to be the consonant with the the best identification rates, and (3) labials /f/ and /v/ are expected to have a better correct identification rate than [+anterior] coronal /z/.

4 Romanian perception experiment

In this section, we investigate the perception of the plain/palatalized contrast with the same fricatives whose acoustic properties were examined in the previous section.

4.1 *Experimental design*

4.1.1 *Subjects and procedure*

Thirty-four subjects took part in this experiment and were compensated for their participation. One of the subjects did not complete the experiment due to technical errors (the program stopped working) so his data were not used in the analysis, which leaves a total of thirty-three subjects (7 males, 26 females), ranging in age from 18 to 30 (mean age = 22.8). All the subjects were native speakers of the standard dialect of Romanian and were tested in Romania.

For the experimental session, the stimuli (targets and fillers) were organized into two 240-trial blocks and presented using E-Prime software for experiment generation with an interstimulus interval (fixation) of 250 ms. The order of the stimuli in each block was randomized for each subject. Listeners were prompted visually with a + sign on the computer screen, which was followed by the audio stimulus. Following the presentation of the stimulus, two words were displayed on the screen, one in the singular form (ending in a plain consonant, e.g. 'pantof') and one in the plural form (phonetically ending in a palatalized consonant, and spelled with a final Ci sequence, according to the orthographic conventions of Romanian, e.g. 'pantofi'). In the case of the fillers, two written words were similarly displayed on the screen, one of which was the actual word that had been read, and the other one a near-homophonous form, e.g. for the word *boi* [boj] 'oxen', the choices were *boi* and *voi* [voj] 'you-pl.'. The position (left or right) of the two words on the screen was randomized. A key was assigned to each of the two positions. Listeners were instructed to press the key corresponding to the target word just heard as rapidly as possible. They were told to make a choice even if they were in doubt, and if no key was pressed within 4000 ms, the next stimulus was presented. There was a short break between the blocks. The program recorded the identification responses and reaction times.

4.1.2 *Stimuli*

The target sentences used in this experiment were selected from those recorded by the subjects in the production experiment. As a general rule, the second utterance of each sentence was used for the perception experiment, unless for some reason this was not possible. The third repetition was used in this case, and if neither the second nor the third repetition were appropriate then the first repetition was used. The reason the selection was made in this order is because there were more hesitations present during the first block due to the novelty of the task, and by the time they recorded the third block the subjects were more tired. The best quality recordings were deemed to be those from the second block.

Sixteen of the thirty-one speakers who recorded the sentences for the production experiment were randomly selected as the speakers for this experiment⁶. It would have been impractical to use all of the speakers, as the number of target items to be included in the perception experiment would have been too high, and consequently the experiment would have been too long overall. The total number of target sentences for this experiment was 5 consonants x 4 items per consonant x 2 forms per item (plain and palatalized) x 16 speakers = 640. Four lists were constructed as follows: the total number of target sentences from the 16 selected speakers (640 items) was divided into 4 parts containing 160 targets each, such that each speaker provided one plain and one palatalized item from each consonant group (the plain and palatalized items produced by the same speaker were counterbalanced, i.e. a target word such as *firav* 'weak' was not uttered by the same speaker in both the singular and plural form within the same list). 4 different subjects were thus needed to cover the entire number of target words (40). Since 16 speakers were included, there were 4 repetitions of each target word. Thus, each of the four lists contained 160 target items. Twice as many fillers (320 sentences) were used in this experiment in order to distract the subjects from the nature of the task. The fillers were also selected among those recorded for the production experiment. Each filler was repeated 4 times, to match the 4 repetitions of any of the target words. The filler list was the same for each of the target lists, and each of the 16 speakers provided 20 different fillers. There were 480 words in total (160 targets plus 320 fillers) in each list. Only one of the lists was presented to any one subject in the perception experiment.

4.2 Data analysis

The dependent variables *accuracy*, *reaction time*, and *sensitivity* were investigated in a repeated-measures within subjects ANOVA with place of articulation (Place) and palatalization status (Pal) as the independent variables.

⁶ The acoustic data from just the sixteen speakers used in the perception study were very similar to the data from all thirty-one original speakers. The table below shows the LDA correct identification for ONLY the subjects who were used as speakers in the perception experiment (compare to Table 1):

	PLAIN	PALATALIZED
f	96.80	69.14
v	88.29	81.38
z	67.02	73.40
ʃ	22.34	94.68
h	100	97.29

Collapsing /f/ and /v/ into a single category, eliminating the laryngeal distinctions, is a practice that can be found in similar studies, for instance the acoustic study of fricatives conducted by Jongman et al. (2000). In his dissertation, Kochetov (2002) also ignored the laryngeal distinction in the analysis of phonotactic patterns of palatalization since, according to his findings, the two classes do not differ with respect to palatalization.

While no further explanation is needed for accuracy and reaction time, the third measure used, sensitivity, requires some comment. This statistical test is also known as the d' (d prime) statistic in the context of the Signal Detection Theory (cf. Wickens 2002). This measure takes bias into account by using both the number of correct responses (how many times a signal was correctly identified; in our case, how many of the palatalized targets were perceived as palatalized), and the number of false alarms (that is, how many times a signal was incorrectly identified; in our case, how many of the plain targets were identified as being palatalized). Based on these, sensitivity (d') scores were computed for each speaker, and the mean sensitivity values for the different places of articulation, as well as for the different consonants, were compared.

4.3 Results

Table 3 provides a summary of the means and standard deviations in terms of accuracy (mean correct identification) and reaction time for each consonant and plain/palatalized condition, as well as the d' values for each consonant. These results include all responses (correct, and incorrect). In the Reaction Time subsection (4.3.2), only the reaction time for correct answers is analyzed.

4.3.1 Accuracy

The results of the repeated measures ANOVA reveal a significant main effect of Place, $F(3, 96) = 347.09$, $p < .001$, as well as a significant main effect of Pal, $F(1, 32) = 4.97$, $p < .05$. The interaction between Place and Pal was not significant. A post-hoc pairwise comparison for Place (with the Bonferroni correction for multiple comparisons) showed that the significant differences in accuracy means were between labial and both [+anterior] and [-anterior] coronals, but not between labial and dorsal. Labials and dorsals were both significantly better identified than coronals, but not from each other. While the identification rates for the two coronal consonants were significantly lower as compared to the other places of articulation, they were also significantly different with respect to each other, with the [+anterior] being correctly identified significantly more often than the [-anterior] coronal.

		ACCURACY (SD)	RT (SD)	d' (SD)
f	plain	91 (29)	1066 (458)	4.4 (1.3)
	palatalized	96 (20)	1045 (429)	
v	plain	94 (24)	1111 (490)	5 (1.4)
	palatalized	98 (13)	1062 (416)	
z	plain	90 (30)	1164 (554)	3.7 (1.4)
	palatalized	93 (25)	1105 (521)	
ʃ	plain	58 (49)	1210 (594)	0.8 (0.7)
	palatalized	65 (48)	1194 (590)	
h	plain	95 (22)	1110 (469)	5.2 (1.5)
	palatalized	98 (14)	1053 (452)	

Table 3

Accuracy (% correct), reaction times (RT) in ms, and d' values for each segment. For Accuracy and RT, N = 528; For d', N = 33. The numbers provided in parentheses represent standard deviations.

Fig. 4. Mean reaction times for correct answers only.

4.3.2 Reaction Time

Table 3 shows reaction times for all answers and, as can be seen, there appears to be a correlation between lower accuracy rates and longer reaction times. We focus now only on the correct answers, shown in Figure 4 for plain and palatalized consonants separately.

Figure 4 shows few significant differences among the different consonant

groups. Once the different places are considered, with the plain and palatalized segments collapsed, the pattern that emerges is strengthened by statistically significant differences. Thus, an ANOVA with Place and Pal as independent variables showed that the factor Place had a significant main effect on reaction times, $F(3,96) = 15.42$, $p < .001$, while Pal and the interaction of these two factors did not. Post-hoc pairwise comparisons with the Bonferroni correction showed that the [-anterior] coronal differed significantly from all other places, and the labial place differed significantly from the [+anterior] coronal.

4.3.3 Sensitivity

First, a repeated measures, within factors ANOVA with d' *sensitivity* scores as the dependent variable and Place (labial, [+anterior] coronal, [-anterior] coronal, and dorsal) as an independent variable⁷, showed that the factor Place had a significant main effect on sensitivity. Post-hoc pairwise comparisons with the Bonferroni correction revealed which pairs were significantly different from each other: dorsal from all other places, [-anterior] coronal from all other places, and labial and [+anterior] coronal from both [-anterior] coronal and dorsal, but not from each other.

A similar analysis was performed with all consonants analyzed separately, such that consonant (Con) was the independent variable and the d' *sensitivity* was the dependent variable. Con was found to have a significant main effect on sensitivity. Post-hoc pairwise comparisons revealed which pairs were significantly different from each other. /j/ was the only segment that differed significantly from all others; the only other significant difference was between /v/ and /z/. Figure 5 displays the sensitivity values for all consonants.

4.4 Summary

We have seen in this section that the results of the perception experiment paralleled very closely the results of the production experiment. All three measures we examined (i.e. accuracy, reaction time, and sensitivity) revealed the same patterns: dorsals tended to be the most favorable hosts for the palatalization contrast, with higher accuracy and sensitivity values. The [-anterior] coronal place of articulation was, by contrast, the least favorable, and differed significantly from all other places with respect to accuracy and reaction time, which were lower, and longer, respectively. Sensitivity values showed that the high accuracy obtained for palatalized [-anterior] coronals cannot be attributed to correct identification, but rather to a strong bias to choose the

⁷ Since accuracy rates for both plain and palatalized segments are used in the computation of the d' statistic, Pal is not a factor here.

Fig. 5. Mean sensitivity by consonant.

'palatalized' response in the case of /ʃ/. The d' value was close to 0, signaling that there was no reliable discrimination of the palatalization contrast with this consonant. As for the comparison between labials and [+anterior] coronals, both accuracy rates and reaction times show that the former were identified significantly better and faster than the latter.

The predictions formulated based on the acoustic results were thus borne out. That is, the three main patterns revealed by the acoustic analysis were also encountered in human perception: (1) sensitivity to the plain-palatalized contrast with post-alveolar /ʃ/ was the lowest, and listeners had a bias for the palatalized form; (2) sensitivity to dorsal /h/ tended to be higher than with all other consonants, and (3) sensitivity to labials was generally higher than with either of the coronal consonants.

5 Discussion

The results reported in the previous sections have enabled the formulation of several robust generalizations with respect to the behavior of secondary palatalization with fricatives in Romanian. These generalizations are validated by the large number of subjects employed in the experiments. Furthermore, the validity of the results is increased by the fact that the same physical stimuli were used in both the production and the perception study.

With regard to the acoustic analysis, it was found that a Bark cepstral analysis yielded very good results not only in terms of its correct classification of plain and palatalized consonants at different places of articulation, but also in its similarity in performance to human listeners. The results are all the more noteworthy if we consider the fact that, except for duration, cepstral coeffi-

cients were the only type of information used. Moreover, an LDA using only c3 and c4 had an accuracy rate that was almost as high as that of the discriminant function using *all* coefficients plus duration. These results support the Bark cepstral analysis as a particularly reliable method for phonetic research.

5.1 *Main patterns*

In terms of production, it was found that the most numerous and consistent acoustic differences between a plain and a palatalized consonant were in the dorsal fricative. A second finding was that none of the acoustic measures differed significantly between the plain and the palatalized form in the case of the post-alveolar fricative. As for the specific comparison between labial and [+anterior] coronal, we found that the labials, both voiced and voiceless, were generally characterized by larger acoustic differences, and a higher number of significant differences between plain and palatalized segments than [+anterior] coronal /z/.

Moving on to the perceptual findings, robust parallels were found between these and the acoustic generalizations presented above. Dorsals had the highest rates of correct identification, followed by the two labials. The coronal segments had lower accuracy and sensitivity rates and of these, the post-alveolar had the lowest of all, suggesting that the listeners are unable to distinguish between the plain and palatalized form. The particularly low salience of the palatalization contrast at this place might be due to the fact that the articulation of /ʃ/ involves a palatal tongue body gesture, and thus the primary and secondary gesture are very similar acoustically. Gestural timing (Zsiga 2000) might also play a role here, with the two gestures overlapping, and thus possibly obscuring, each other. There is one other possibility, namely the complete absence of the secondary gesture, but this can be dismissed in light of a more in-depth analysis of the production data (Spinu 2010). While no significant acoustic differences between the plain and palatalized form were found at the group level, a different picture emerged when the subjects were considered individually, with only 4 out of the 31 speakers examined failing to produce any significant difference between plain and palatalized /ʃ/.

In sum, the patterns we uncovered do not replicate previous cross-linguistic findings with stops. This provides a negative answer to the question whether the behavior of palatalization is uniform across manners of articulation (Kochetov 2002, p. 51). The next question to be addressed is why exactly this is so - specifically, what is it about the stop manner of articulation that favors secondary palatalization in coronals, while this is not the case with fricatives? In the next sections, the nature and implications of these differences are examined.

5.2 *Realization of secondary palatalization*

We start with a closer look to the realization of each type of segment. As discussed in Section 3.1.2, the only difference between the plain and palatalized form of the fricatives we investigated is in the absence versus presence of the additional palatal gesture, without other changes in place or manner of articulation, as is the case with stops.

As for palatalized stops, a comprehensive description of their acoustic properties is provided in Kochetov 1999. The author discusses aspects such as the formant release (CV transitions) and approach (VC transitions) and the burst of palatalized consonants. With respect to the burst, according to Bolla (1981) and Halle (1959), among others (cf. Kochetov 1999), one of the acoustic parameters of Russian sounds is the burst of fricative noise that follows the closure of a stop. As far as the secondary palatalization contrast is concerned, in addition to overall duration (with palatalized segments being longer than plain segments), palatalized segments also differ in the duration of the burst: 5 ms for [p] and 20 ms for [pʲ] (based on Bolla 1981, cf. Kochetov 1999). The burst duration not only enhances the contrasts between plain and palatalized stops, but also serves to distinguish the different places of articulation of palatalized stops. While the bursts of palatalized labial and velar stops are 20 ms and 55 ms respectively, the burst of the coronal stop is 98 ms (Bolla 1981), which is close to the acoustic structure of an affricate. Because of this, the palatalized coronal stop emerges as most salient. As for the quality of the burst, in Standard Russian the palatalized /tʲ/ is also accompanied by affrication, that is, by the palatalized element [sʲ].

To summarize, in the case of fricatives, the strength of the distinction between plain and palatalized segments in labials and dorsals suggests that palatalization, as a secondary gesture, is more salient when the primary place gesture takes place farther away from the palate and/or when there are distinct articulators for the two gestures involved: lips or glottis for the primary place, and tongue body for the secondary place. By contrast, the reverse holds true for stops, with the cases where the primary and secondary articulation are close to each other being most salient. The main difference between the realization of secondary palatalization in fricatives versus stops is that in the latter, secondary palatalization is accompanied by a release burst followed by strong affrication. This pattern can be accounted for from the perspective of Enhancement Theory.

5.3 *Enhancement Theory*

Enhancement Theory (Stevens et al. 1986, Stevens and Keyser 1989, 2010) seeks to explain certain patterns of regular cross-linguistic variation. Starting

from the fact that languages tend to preserve useful contrasts, it proposes that supplementary features and gestures may be employed to reinforce existing contrasts between two sounds or sound sequences along an acoustic dimension that distinguishes them. After being introduced in a linguistic system, these features tend to survive, and may eventually come to replace the feature which they originally served to enhance. For instance, in some Irish dialects velarization has come to replace secondary palatalization, such that phonemically plain consonants are now velarized and phonemically palatalized ones are plain (Kochetov 2002). Hayes and Steriade (2004) propose that cross-linguistically, perceptually fragile contrasts (due to relative lack of perceptibility) tend to undergo one of two changes: they are either enhanced, or neutralized.

As far as secondary palatalization is concerned, one possibility is that enhancement strategies are often used with coronals, but not (to the same extent) with labials and dorsals. We have seen that in the fricatives we investigated, coronals were at a disadvantage as far as the acoustic and perceptual properties of the plain-palatalized contrast are concerned. Crucially, the addition/strengthening of frication following the release burst which is characteristic of palatalized coronal stops was not (could not be) implemented with these fricatives. While palatalized coronal stops can also be produced without these cues (e.g. [tʲ] in Romanian, [peftʲ] 'fish-pl.']), it has been noted that there is a strong cross-linguistic tendency for /tʲ/ to have a wide range of realizations, among which are the affricates [ts], [tʃ] and [tɕ] (Kochetov 1998, 2002). Furthermore, Mester and Itô note that, while the characterization 'palatalized' is strictly speaking only accurate for noncoronals, i.e. labials and velars, palatalization of coronals (*t, d, s, z, n*) changes their primary place of articulation to palatal/alveopalatal⁸ (Mester and Itô 1989). In light of Enhancement Theory, processes such as affrication, assibilation, or the change in primary place of articulation that are known to accompany secondary palatalization at the coronal place might be conceived of as having arisen in a language for the purpose of enhancement.

If this is indeed the case, then we must be careful in interpreting the place of articulation asymmetry proposed by Kochetov (Kochetov 2002) and based, among other factors, on the perceptual properties of plain and palatalized labial and coronal stops of Russian. A distinction must be made between enhanced secondary palatalization, and secondary palatalization that is not accompanied by enhancement, manifested purely as articulatory overlap of two place gestures. It is undoubtedly the case that the plain-palatalized contrast with coronals is favored across languages in many relevant respects: it appears in more grammars, it is more common, and less likely to be neutralized. If the privileged status of coronals, however, is due to the various enhancement

⁸ Note that while this may be a strong tendency, it is not necessarily the case, for example in Romanian we encounter palatalized forms of [t], [n], and [z] without an additional change in primary place.

strategies used in languages to reinforce relatively fragile contrasts, then the secondary palatalization contrast with coronals should in fact be regarded as the most perceptually fragile when it comes to secondary palatalization.

If coronals are not favorable to the realization of the plain-palatalized contrast unless they are enhanced, the question also arises why they are so seldom the target of neutralization (and why neutralization of the contrast in a language targets the labials first, in an implicational fashion). It has been suggested that inventories can be influenced by factors other than perceptual salience, for instance ease of articulation and featural economy (Ladefoged 2001, Hayes and Steriade 2004). Bateman (2007) proposes that, with respect to secondary palatalization, the less unmarked case is for both the primary and secondary articulator to be lingual, while having two different articulations (e.g. labial and lingual) is more marked. This account successfully explains why labial secondary palatalization in a given language is dependent on either coronal or dorsal palatalization (in an implicational fashion), and why it is a preferred target of neutralization.

To conclude, it may well be the case that, in the absence of additional enhancement strategies, the palatalization contrast is better realized acoustically at the labial and dorsal place (versus coronal), as suggested by the salience scale that emerged in our study. For reasons of availability of enhancement strategies, gestural coordination (employ less rather than more articulators) or some type of featural economy, however, it is the palatalized coronals that are preferred in phonemic inventories.

6 Conclusions

We provided a detailed account of the acoustic and perceptual properties of secondary palatalization in Romanian fricatives. We found that in fricatives the labial and dorsal places of articulation exhibit more salient effects of secondary palatalization than the coronal place. This is different from secondary palatalization in stops, which is frequently characterized by the presence of additional enhancement strategies at this place of articulation. Since this type of enhancement consists of additional frication, fricatives at the coronal place of articulation do not benefit from it. This being the case, the place of articulation markedness hierarchy proposed by Kochetov (2002) with respect to secondary palatalization may need rethinking. While palatalization is more frequently encountered with coronals cross-linguistically, this may be due to its special phonological status as an enhanced place of articulation, as opposed to an unmarked one.

Acknowledgments

Work partially supported by NSF doctoral dissertation research improvement grant # 0720231.

References

- [1] Bateman, N. 2007. *A crosslinguistic investigation of palatalization*. PhD dissertation, Department of Linguistics, University of California, San Diego.
- [2] Bhat, D.N.S. 1978. A general study of palatalization. In Greenberg, J.H. (ed.), *Universals of human language, vol. 2: phonology*. Stanford, CA: Stanford University Press.
- [3] Bolla, K. 1981. *A Conspectus of Russian Speech Sounds*. Budapest: Hungarian Academy of Science.
- [4] Bunnell, H.T., Polikoff, J.B. and McNicholas, J.E. 2004. Spectral Moment vs. Bark Cepstral Analysis of Children’s Word-initial Voiceless Stops. *Proceedings of the 8th International Conference on Spoken Language Processing*. Jeju, Korea.
- [5] Chitoran, I. 2002. *The Phonology of Romanian: A Constraint-based Approach*. Berlin & New York: Mouton de Gruyter.
- [6] Clements, G. N. and Hume, E. 1995. Internal Organization of Speech Sounds. In *The Handbook of Phonological Theory* (published 1996).
- [7] Comrie, B. 1981. *The languages of the Soviet Union*. Cambridge: Cambridge University Press.
- [8] Fant, G. 1970. *Acoustic theory of speech production: With calculations based on x-ray studies of Russian articulations*. Second Edition (Printing). The Hague: Mouton.
- [9] Ferragne, E. and Pellegrino, F. 2010. Vowel systems and accent similarity in the British Isles: Exploiting multidimensional acoustic distances in phonetics. *Journal of Phonetics* 38, 526–53.
- [10] Hayes, B. and Steriade, D. 2004. Introduction: The phonetic bases of phonological markedness. In Hayes, B., Kirchner, R. and Steriade, D. (eds.), *Phonetically-based phonology*. Cambridge University Press.
- [11] Hume, E. 1994. *Front Vowels, Coronal Consonants and Their Interaction in Nonlinear Phonology*. Garland Publishing, Inc.
- [12] InvTool. 2007. Speech Inventory Construction Tool (Version 3.2.13.6). [Computer Program]: The Nemours Foundation.

- [13] Iscrulescu, C. 2003. Morphological Faithfulness and Phonological Markedness: The Case of Romanian Nominals. In Nihan Ketrez, F. et al. (eds.), *USC Working Papers in Linguistics 1*. Los Angeles: University of Southern California.
- [14] Jongman, A., Wayland, R. and Wong, S. 2000. Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America* 10/2000; 108(3 Pt 1).
- [15] Kavitskaya, D. 2006. Perceptual salience and palatalization in Russian. In Goldstein, L., Whalen, D. H. and Best, C.T. (eds.), *Laboratory phonology 8*. Berlin, New York: Mouton de Gruyter.
- [16] Kenstowicz, M. 1972. Lithuanian phonology. *Studies in the linguistic sciences*, 2:1-85. Urbana, IL: University of Illinois.
- [17] Kochetov, A. 1998. Articulatory Gestures and Perceptual Enhancement: Palatalized Labials in Polish Dialects. *Proceedings of the First High Desert Student Conference in Linguistics*. Albuquerque, NM: University of New Mexico.
- [18] Kochetov, A. 1999. Phonotactic constraints on the distribution of palatalized consonants. *Toronto Working Papers in Linguistics*, Vol 17.
- [19] Kochetov, A. 2002. *Production, Perception, and Emergent Phonotactic Patterns: A Case of Contrastive Palatalization*. New York: Routledge.
- [20] Ladefoged, P. 1993. *A course in phonetics*, 3rd edition. San Diego: Harcourt Brace Jovanovich.
- [21] Ladefoged, P. 2001. *Vowels and consonants*. Oxford: Blackwell.
- [22] Mallinson, G. 1986. *Rumanian*. Croom Helm: London.
- [23] Mester, R. A., and Itô, J. 1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65. 258-293.
- [24] Ní Chiosáin, M. 1994. Irish palatalisation and the representation of place features. *Phonology* 11, 89-106.
- [25] Padgett, J. 2001. Contrast Dispersion and Russian Palatalization. In Hume, E. and Johnson, K. (eds.), *The Role of Speech Perception in Phonology*. New York: Academic Press.
- [26] Padgett, J. 2003. The emergence of contrastive palatalization in Russian. In Holt, E. (ed.), *Optimality Theory and language change*. Kluwer Academic Press, Dordrecht. 307-335.
- [27] Ruhlen, M. 1973. *Rumanian Phonology*. PhD dissertation, Stanford University.
- [28] Spinu, L. 2007. Perceptual Properties of Palatalization in Romanian. In Camacho, J. et al. (eds.), *Romance Linguistics 2006: Selected papers from the 36th Linguistic Symposium on Romance Languages (LSRL)*, New Brunswick, March-April 2006. Amsterdam & Philadelphia: John Benjamins.

- [29] Spinu, L. 2009. Romanian palatalization: the role of place of articulation in perception. In Masullo, P. J., O'Rourke, E., Huang, C-H. (eds.), *Romance Linguistics 2007*. Amsterdam & Philadelphia: John Benjamins.
- [30] Spinu, L. 2010. *Palatalization in Romanian: Experimental and Theoretical Approaches*, doctoral dissertation, University of Delaware.
- [31] Stevens, K., Keyser, S.J. and Kawasaki, H. 1986. Toward a phonetic and phonological investigation of redundant features. In Perkell, J. and Klatt, D.H. (eds.), *Symposium on Invariance and Variability of Speech Processes*, Hillsdale, NJ: Lawrence Erlbaum, Ch. 20, 426-463.
- [32] Stevens, K. and Keyser, S. 1989. Primary features and their enhancement in consonants. *Language* 65.1, 81-106.
- [33] Stevens, K. and Keyser, S. 2010. Quantal theory, enhancement and overlap *Journal of Phonetics* 38, Issue 1, 10-19. Stevens, K.N.; Keyser, S.J.
- [34] Vogel, I. and Spinu, L. 2009. The Domain of Palatalization in Romanian. In Masullo, P. J., O'Rourke, E., Huang, C-H. (eds.), *Romance Linguistics 2007*. Amsterdam & Philadelphia: John Benjamins.
- [35] Wickens, T. 2002. *Elementary Signal Detection Theory*. New York: Oxford University Press.
- [36] Zsiga, E. 2000. Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28, 69-102.

Appendix

Target words used in both the acoustic and perceptual experiments.

Words ending in	Singular	Plural	Translation
f	panto[f]	panto[f ^j]	Shoe/s
	vata[f]	vata[f ^j]	Bailiff/s
	carto[f]	carto[f ^j]	Potato/es
	zulu[f]	zulu[f ^j]	Curl/s
v	bolna[v]	bolna[v ^j]	Sick/pl
	groza[v]	groza[v ^j]	Great/pl
	fira[v]	fira[v ^j]	Feeble/pl
	zugra[v]	zugra[v ^j]	House painter/s
z	obe[z]	obe[z ^j]	Obese/pl
	ursu[z]	ursu[z ^j]	Morose/pl
	moflu[z]	moflu[z ^j]	Grumpy/pl
	chine[z]	chine[z ^j]	Chinese/pl
ɟ	coda[ɟ]	coda[ɟ ^j]	Slacker/s
	cire[ɟ]	cire[ɟ ^j]	Cherry tree/s
	coco[ɟ]	coco[ɟ ^j]	Rooster/s
	ghidu[ɟ]	ghidu[ɟ ^j]	Playful/pl
h	caza[x]	caza[x ^j]	Cossack/s
	vala[x]	vala[x ^j]	Wallachian/s
	mona[x]	mona[x ^j]	Monk/s
	paro[x]	paro[x ^j]	Vicar/s

Fig. 6. Spectrograms of words ending in plain-palatalized pairs. In all cases, the consonant of interest is the final fricative segment. The spectrograms were obtained using the Praat program for acoustic analysis. The plain fricatives are on the right, and the palatalized ones on the left. Palatalization is denoted with a word-final 'j', and post-alveolar /ʃ/ is transcribed with an 'S.'