Interspeech 2022
18-22 September 2022, Incheon, Korea

Voicing neutralization in Romanian fricatives across different speech styles

Laura Spinu[1], *Ioana Vasilescu*[2], *Lori Lamel*[2], *Jason Lilley*[3]

[1]City University of New York, Kingsborough Community College
[2]Interdisciplinary Laboratory of Digital Sciences, Paris-Saclay University
[3]Nemours Biomedical Research

lspinu@kbcc.cuny.edu, ioana@limsi.fr, lamel@limsi.fr, lilley@asel.udel.edu

Abstract

Romance languages such as Italian or Spanish preserve fricative voicing contrasts in word-final position, while their neutralization has been reported for European Portuguese, but the behavior of Romanian fricatives remains understudied. Previous work with Romanian fricatives suggests a pattern of final devoicing but, due to the specific properties of the corpus analyzed, it is unclear if this is limited to the presence of secondary palatalization and/or the result of morphological conditioning. In this study, we apply speech processing tools to investigate the acoustic characteristics of the voicing contrast in fricatives in contemporary spoken Romanian. We examine a corpus of prepared speech from newscasts and semi-spontaneous TV debates (86 speakers) and compare our results to previous findings from a corpus of controlled experimental speech (31 speakers). Our classification tool employs cepstral coefficients and hidden Markov model (HMM)-defined temporal regions to identify the properties of these segments. Our findings conform to typological predictions regarding partial devoicing in coda position, especially at more posterior places, but we find little support for voicing neutralization in Romanian fricatives more generally. Our study thus documents the properties of Romanian fricatives and contributes to our understanding of the dynamics of contrast maintenance in phonological systems.

Index Terms: Romanian fricatives, voicing contrast, devoicing, neutralization, speech production, cepstral coefficients

1. Introduction

The production of voiced fricatives involves "the complex interaction of articulatory constraints from three separate goals: the formation of the appropriate oral constriction, the control of airflow through the constriction so as to achieve frication, and the maintenance of glottal oscillation by attending to transglottal pressure") [1]. The extent to which this articulatory complexity plays a part in fricative devoicing varies from language to language. While fricative devoicing has not been extensively reported for other Romance languages [2], studies of European Portuguese [3, 4, 5] indicate that heavy devoicing occurs in coda position (up to 76.5% of the time full devoicing for [z] and 48.4% for [v], with additional partially devoiced segments).

Given Romania's status as an island of "latinity" [6] geographically surrounded for the most part by Slavic languages, as well as the presence of a number of minority languages (e.g. Ukrainian, Serbian, Croatian, Bulgarian, Slovak, Turkish, German, and Russian), all of which exhibit final obstruent devoicing, the question arises if Romanian has evolved to be more similar to Portuguese as a result of language contact phenomena. Other features of Slavic phonology have been documented in Romanian, for example two- or three-member onset consonant clusters such as [vl], [zm], [zdr], which do not exist elsewhere in the Romance family and were not attested at earlier stages of Romanian either, but are thought to have appeared in this language through the borrowing of Slavic terms [7]. An earlier study of Romanian reported partial devoicing of the labiodental fricative contrast in word-final position [8], but the corpus employed contained in equal parts plain and secondarily palatalized fricatives corresponding exclusively to either singular (plain) or plural (palatalized) forms of nouns and adjectives, as secondary palatalization in Romanian is for the most part morphologically conditioned. Therefore, the possibility arises that final devoicing is associated with either the presence of secondary palatalization or that of a morphological marker in this language. To verify this possibility, environments outside of the word-final position, as well as the plural morpheme, should be examined. Our goal is to do so using a relatively novel metric which has been increasingly employed in the field of acoustic phonetics over the past decade, specifically cepstral coefficients.

Since they were 'borrowed' from the speech processing literature into linguistic studies, cepstral coefficients have proven to be an informative measure regarding the strength of various types of contrasts [9], including the voicing contrast in fricatives. Thus, cepstral coefficients were previously successful in categorizing English obstruent bursts [10], English vowels [11], Romanian fricatives [12, 13], Russian sibilant fricatives [14], Azerbaijani fricatives [15], and Greek fricatives [16]. The potential advantages of using this measure in phonetic studies have been discussed extensively in recent literature. For instance, [11] recommended Mel-frequency cepstral coefficients (MFCCs) as a means to compute distances between vowels. This method yielded a very good estimate of the acoustic distance between 13 different accents of the British Isles, leading the authors to conclude that "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" (p. 536). [17] obtained 85% correct classification for three places of articulation in English fricatives from the TIMIT corpus using a set of 13 MFCCs, suggesting that the advantages of cepstral coefficients remain robust despite dialectal variation (in this case, 8 dialects of English) and large samples of speakers (that is, 630 speakers in this corpus).

While many of the studies listed above focused on the classification of place of articulation, other studies have focused on, or also included, voicing classification. MFCCs were used successfully in a study on the classification of voicing in fricatives in British English and European Portuguese [18]. Among other analyses conducted on a corpus of 1,522 intervocalic Greek fricatives produced by 29 monolingual speakers, [16] classified voicing in two pairs of front fricatives, labiodentals and interdentals, using Bark-scaled cepstral coefficients. The measures employed in this study were as follows: 18 measures obtained from the preceding vowel (6 cepstral coefficients $\times$ 3 tempo-

  10.21437/Interspeech.2022-10716

ral regions), 18 measures from the fricative, and 6 measures from the first region of the vowel following the fricative. To these were added the onset and offset transition feature vectors (totaling 12 measures), producing a total of 54 measures per target fricative. The overall classification accuracy for voicing was very high (95.2% based on all the predictors extracted, and 83.9% based on only the top 5 predictors), indicating that voicing is robustly encoded in each fricative in Greek.

2. Previous findings with Romanian

[8] compared the performance of cepstral coefficients and spectral moments in a study classifying a corpus of 3,674 word-final Romanian fricatives produced by 31 native speakers by place (four places of articulation), voicing, secondary palatalization, and gender, and found that the former method yielded higher classification rates across the board, regardless of whether these measures came from HMM-defined regions or equal-duration regions inside a segment. When the classification of voicing in the [f-v] pair was based on acoustic measures from the onset of frication noise the correct classification rate was above 95%. By contrast, when the acoustic information was obtained from the middle and the end regions of a fricative, correct classification rates decreased to below 80% and 57%, respectively. Given that only two consonants are classified, this is only slightly higher than chance. This suggests that the voicing distinction is mostly realized at the beginning of the fricative, and is consistent with the possibility that Romanian fricatives tend to devoice towards the end of the segment. A similar decrease in classification accuracy was obtained by classifying voicing in the larger corpus containing five fricatives at four places of articulation (i.e. [f, v, z, ʃ, h]), with a notable difference: the last region of the fricatives yielded approximately 70% accuracy (see Figure 1). Note that the other places of articulation were not considered separately in this study because of the lack of voicing contrasts.

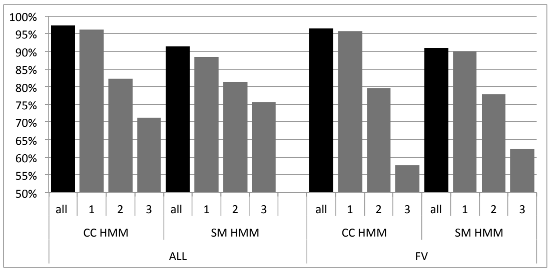

Figure 1: *Voicing classification results in [8] for the cepstral coefficient (CC HMM) and spectral moment (SM HMM) parameter sets for logistic regression models trained either on all parameters (black column), or just parameters extracted from a single region (gray columns). Left: models trained on all 5 consonants. Right: models trained on the labial subset.*

The results presented in Figure 1 suggest the occurrence of a phenomenon of partial devoicing, reflected in the increased similarity of acoustic cues that make for lesser category separation. This hypothesis was corroborated by reconstructed spectra averaged over all speakers and repetitions indicating that the voice bar for the voiced segments /v, z/ was almost entirely absent in the final two regions. It remains unclear, however, to what extent this is a place of articulation effect, a speech register effect (as this was a corpus of controlled, laboratory-elicited speech), or whether the word-final position and the presence of the plural morpheme and/or secondary palatalization played an important part towards the devoicing. The goal of the present study is to clarify this issue.

3. Method

In this study, we employ cepstral measures to further explore the voicing contrast in Romanian fricatives. We expand our investigation to a new speech register, specifically a corpus of semi-prepared and spontaneous speech [19, 20]. Compared to previous work, this new corpus includes more speakers and additional phonotactic environments in word-initial, word-medial, and word-final position. Our goal is to determine whether Romanian fricatives exhibit neutralization of the voicing contrast, as manifested by full or partial devoicing of these segments. Our research question is whether the previously observed pattern of devoicing [8] is restricted to the specific word-final position, the plural morphological context, and/or labial place of articulation, or more generally manifested in the language.

3.1. The corpus

To address our research question, we employ a corpus of broadcast data collected from various Romanian radio and television shows (7 hours and 10 minutes). This corpus includes four different sources, comprising two speaking styles: semi-prepared, read speech and more spontaneous interactions such as debates [19, 20]. The number of speakers varies according to the source, ranging from 3 to 24 and includes 86 males and females in total, all speakers of the standard version of Romanian based on the Southern dialect. Efforts were made to avoid sources with significant quantities of overlapping speech, foreign or regional accents and noisy backgrounds. The high-quality data (16kHz sampling rate) were collected and annotated in the context of the Quaero program[1].

More specifically, the data were aligned and segmented using the system described in [19], for which they served as development and evaluation corpora. Forced alignment mode was used. Since the manual transcription was provided *a priori*, the system only had to select the best matching pronunciation for each word, and the corresponding locations of word and phone boundaries. The pronunciation lexicon and acoustic models are based on a set of 29 phones, including 20 consonants, 7 vowels and 2 glides. The acoustic models were trained in a semi-supervised manner using approximately 400 hours of unannotated audio [19]. Due to acoustic modeling constraints, phone segments have a minimum duration of 30 ms (3 frames). The total number of words identified is 56,296, of which 9,032 were distinct words.

3.2. The target segments

The target segments we analyzed included pairs of voiced and voiceless fricatives from three places of articulation: labiodental /f, v/, dental /s, z/ and postalveolar /ʃ, ʒ/. We extracted these segments from three positions inside words: initial, medial[2], and final. Given the fact that we use a more naturalistic type of corpus, these segments were not balanced in terms of their frequencies (see Table 1), nor in terms of their preceding/following segments. It should be noted that only 1.37% of

[1] http://www.quaero.org

[2] We did not obtain information regarding syllabic position for the word-medial segments and we are thus unable to classify them as onset or coda consonants.

the final segments identified were accompanied by word-final secondary palatalization, compared to 50% in the corpus employed in [8] and therefore our study is not a direct comparison but rather a complementary extension to previous findings.

Table 1: *Relative frequencies of voiced (+V) and voiceless (-V) fricatives in our corpus.*

	Initial		Medial		Final	
	+V	**-V**	**+V**	**-V**	**+V**	**-V**
Dental	19%	81%	23%	77%	11%	89%
Labiodent.	45%	55%	67%	33%	91%	9%
Postalv.	7%	93%	20%	80%	36%	64%
Count	976	4,145	2,037	4,161	89	267

3.3. Analysis

A series of feature vectors comprising six Bark cepstral coefficients (DC and the first five cosine terms) were extracted from the fricatives of interest, using overlapping Hamming analysis windows, 20 ms wide and spaced 10 ms apart. Next, HMMs were used to divide the segments into three temporal regions of internally minimized variance. HMMs partition the time-varying structure of segments into a series of piecewise approximately stationary regions, modeling the dynamic nature of the segments. These separate temporal regions offer the possibility to examine the strength of the acoustic cues to voicing progressively throughout each segment, addressing the possibility that voicing might affect only part of a segment (partial devoicing), specifically the middle and/or final regions [21]. Examining acoustic properties region by region, when the regions are determined based on internal variance, makes it more likely that if at all progressive the acoustic consequences of devoicing might be captured.

After dividing each target segment into regions, the means of the features over all of the vectors in each region were calculated and used as input to the statistical analyses. This resulted in 18 different measures: 6 cepstral coefficients $\times$ 3 regions. Following McMurray and Jongman [22], we used Matlab R2013a [23] to conduct multinomial logistic regression analyses with voicing as the dependent variable and the 18 measures as continuous explanatory variables. Classifications by voicing were performed for each place subset, that is, for the labiodental, dental, and postalveolar consonants separately.

3.4. Results

Figures 2-4 display the mean correct classification by voicing for each of the three places of articulation. A segment counted as correct if its classification matched its phonemic description with respect to its voicing status (voiced or voiceless). Contrary to previous findings, labiodentals yielded very high correct classifications, reaching 100% in all regions of word-final fricatives. In word-initial and word-medial position, classification accuracy was around 90% for the second and third region of the fricative, but approximately 10-13% lower in the initial region. We attribute this decrease to the fact that the first region is more likely to be affected by coarticulation from a preceding segment leading to erroneous classification of the voiceless tokens as voiced. In our corpus, about 75% of all segments preceding the target fricatives were vowels or glides, and only 12% of the preceding environments were voiceless. That this phenomenon is restricted to labiodentals may have to do with their place gesture being independent of the lingual articulator, leading to increased coarticulation with preceding segments whose main place gesture is realized with the tongue.

The situation is somewhat different for the other two places, which both exhibit lower correct classification rates in medial and final position. In the case of the dental fricatives, word-initial segments yield the highest correct classification rates, which only slightly decrease as we move towards the end of the fricative. By contrast, word-medially and word-finally we note that the classification rate increases as we approach the end of the segment, suggesting that the voicing cues are robust throughout the entire fricative. Again, we tentatively attribute the decreased classification rates of the initial regions to coarticulation with preceding segments.

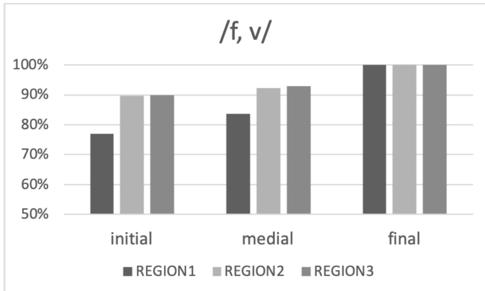

Figure 2: *Voicing classification results for labiodentals.*

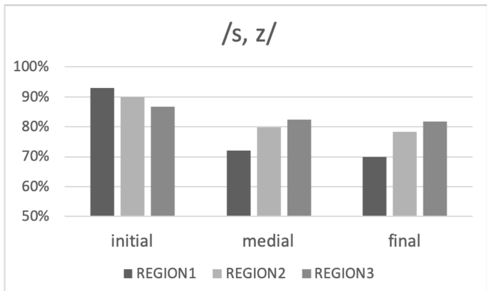

Figure 3: *Voicing classification results for dental fricatives.*

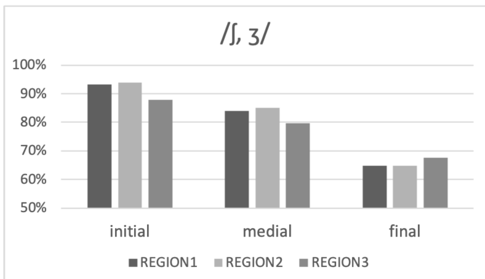

Figure 4: *Voicing classification results for postalveolars.*

Turning to the postalveolar place, we observe much lower correct classification rates in the final position compared to the other two places, ranging from 64.7% in the first region to 67.6% in the third region of the fricative. These numbers suggest that voicing cues are encoded less robustly in word-final position at this place. Just like dentals, the word-initial position appears most favorable for the classification of voicing, with the classification accuracy exceeding 90% for the initial and middle regions, and decreasing only slightly (to 88%) in the final region. The accuracy remains relatively constant throughout all regions of word-medial postalveolar fricatives.

Examination of the error patterns revealed that for the most part a similar amount of errors were made in the classification of voiced and voiceless tokens in each category (that is, voiced fricatives being classified as voiceless and vice versa), typically between 6-20%. This pattern was very consistent in word-initial and word-medial position across all places of articulation, and also in word-final position for labiodentals and dentals. The only asymmetries were noted for postaolveolars in word-initial position, where higher rates of voiced tokens were classified as voiceless (41.2%) than vice versa (29.4% of voiceless tokens were classified as voiced), as well as in word-final position where the reverse pattern was noted: 41.2% of voiceless tokens were classified as voiced, compared to only 23.5% of voiced tokens categorized as voiceless. The reasons behind this asymmetry are not clear, but the possibility arises that differences in the relative numbers of voiced versus voiceless segments intrinsic to the naturalistic nature of our corpus might affect the results. Further investigation of this topic is recommended.

4. Discussion

We investigated whether the devoicing pattern in Romanian labiodental fricatives reported in [8] is indicative of a more general obstruent devoicing pattern in Romanian, as previously attested with another Romance language, Portuguese [4, 5, 21], or whether it might be restricted to the specific context examined in [8] – specifically, that of a word-final plural morpheme accompanied by secondary palatalization. To address this question, we expanded on previous work in several ways: by analyzing a larger number of speakers (86 compared to 31), a novel, more naturalistic speech register (news broadcasts and debates, compared to controlled lab speech), and 3 different environments (word-initial, word-medial, and word-final, compared to only word-final in [8]). We employed a method based on cepstral coefficients and temporal regions, which recent literature has recommended for various types of classification, including voicing, cross-linguistically. Our assumption was that higher correct classification rates for a pair of phonemes at a particular place of articulation would reflect more robust acoustic cues to the voicing contrast at that place.

Our results did not replicate previous findings of final devoicing in labiodentals [8]. In fact, we obtained the highest correct classification rates at this place, with 100% accuracy in word-final position, departing from previous work. However, it should be noted that, given the limited number of tokens in this particular position, further investigation is recommended. We tentatively conclude that the pattern reported in [8] indicates devoicing that is triggered by the presence of a plural morpheme or the secondary palatalization gesture that typically accompanies this morpheme in Romanian [12, 24]. It is worth noting that in [8] it may be the combination of both secondary palatalization and the word-final position that triggers devoicing due to increased articulatory complexity.

While we found no evidence of final devoicing in labiodentals, we did observe partial (about 20%) devoicing in word-medial and word-final dental fricatives and more pronounced (about 35%) devoicing in word-final postalveolar fricatives. Our findings are in line with those of previous studies that have found a consistent increase in devoicing with more posterior place of articulation [25, 26]. Since, unlike in [8] secondary palatalization was only present in 1.37% of our corpus, the most likely explanation is that the partial devoicing we observed is subject to positional effects (note that word-medial segments could also have been in coda position in our corpus). Coda position has long been claimed to be a weaker position for the realization of acoustic cues for consonants in general [27, 28, 29, 30]. It is thus not surprising that coda laryngeal contrasts are typologically more marked than onset laryngeal contrasts. The voiced-voiceless distinction has been neutralized in European Portuguese, Russian, German, Turkish (and many other languages) in this position.

These observations raise questions related to the dynamics of contrast maintenance in phonological systems. Given that simultaneous maintenance of voicing and frication is rather challenging [25], the devoicing patterns we have uncovered are expected typologically, though based on our corpus alone we cannot speculate whether this situation reflects an ongoing process of language change. Without additional data to compare our findings to, it remains unclear whether the partial devoicing we observed with dentals and postalveolars represents a stable synchronic situation (with perhaps only certain speakers partially devoicing due to subtle dialectal differences) or a change in progress affecting most of the speakers to some extent, possibly subject to gender or age effects [32, 31]. Lastly, because our corpus contained both semi-prepared read speech as well as naturalistic debates, the possibility arises that the devoicing we observed is related to temporary aerodynamic constraints imposed by more irregular breathing in spontaneous, animated speech. That this happened more frequently in dentals, and especially in postalveolars, confirms previous generalizations that the voicing contrast at these places of articulation is typologically weaker and more vulnerable to neutralization [25, 26].

5. Conclusions

Our study documents the properties of the voicing contrast in fricatives in an understudied language by considering higher number of speakers, more varied syllabic positions, and a new speech register compared to previous work. Our contribution is noteworthy in its use of a large-scale corpus of spoken Romanian (often referred to as a less-resourced language) produced in various communicative frameworks. This corpus is likely to reflect the trends present in contemporary spoken Romanian, complementing previous work with controlled speech produced in lab settings. Our findings suggest that Romanian behaves similarly to most other Romance languages (except European Portuguese), despite contact with several minority languages which exhibit final obstruent devoicing, including Ukrainian, Turkish, German, and Russian. We have however found evidence of limited partial devoicing in postalveolars and to a lesser extent in dental fricatives. In line with previous work, we have identified the coda position as more vulnerable to neutralization. Both of these trends conform to typological predictions. By comparing our findings to previous work, we have concluded that final devoicing also affects labiodentals when secondary palatalization is present word-finally, though the strength of this conclusion remains dependent on future work with higher numbers of tokens in final position and more balanced datasets.

6. Acknowledgements

Spinu: dissertation grant #0720231 (National Science Foundation); Vasilescu: DATAIA/MSH Paris-Saclay grant “OTELO”.

7. References

[1] M. I. Proctor, C. H. Shadle, and K. Iskarous, "Pharyngeal articulation in the production of voiced and voiceless fricatives," *J. Acoust. Soc. Amer.*, vol. 127, pp. 1507–1518, 2010.

[2] C. Shih, B. Möbius, and B. Narasimhan, "Contextual effects on consonantal voicing profiles: A cross-linguistic study," in *Proc. 14th Int. Congr. Phonetic Sci.*, San Francisco, CA, USA, Aug. 1999, vol. 2, pp. 989–992.

[3] L. M. Jesus and C. H. Shadle, "A parametric study of the spectral characteristics of European Portuguese fricatives," *J. Phonetics*, vol. 30, no. 3, pp. 437–464, 2002.

[4] D. Pape and L. Jesus, "Production and perception of velar stop (de)voicing in European Portuguese and Italian," *EURASIP J. on Audio, Speech, and Music Process.*, vol. 6, pp. 1–10, 2014.

[5] D. Pape and L. Jesus, "Stop and fricative devoicing in European Portuguese, Italian and German," *Lang. and Speech*, vol. 58, no. 2, pp. 224–246, 2015.

[6] A. M. Dudau, "Romanian language and its dialects," *Res. and Sci. Today*, vol. 7, p. 183, 2014.

[7] P. R. Petrucci, *Slavic Features in the History of Rumanian*. LINCOM Europa, 1999.

[8] L. Spinu and J. Lilley, "A comparison of cepstral coefficients and spectral moments in the classification of Romanian fricatives," *J. Phonetics*, vol. 57, pp. 40–58, 2016.

[9] B. Milner and J. Darch, "Robust Acoustic Speech Feature Prediction From Noisy Mel-Frequency Cepstral Coefficients," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 19, no. 2, pp. 338-347, Feb. 2011, doi: 10.1109/TASL.2010.2047811.

[10] H. T. Bunnell, J. B. Polikoff, and J. E. McNicholas, "Spectral moment vs. bark cepstral analysis of children's word-initial voiceless stops," in *Proc. INTERSPEECH 2004 – 5th Annu. Conf. Int. Speech Commun. Assoc.,* Jeju Island, Korea, Oct. 2004, pp. 1313–1316.

[11] E. Ferragne and F. Pellegrino, "Vowel systems and accent similarity in the British Isles: Exploiting multidimensional acoustic distances in phonetics," *J. Phonetics*, vol. 38, no. 4, pp. 526–539, 2010.

[12] L. Spinu, I. Vogel, and T. Bunnell, "Palatalization in Romanian: Acoustic properties and perception," *J. Phonetics*, vol. 40, no. 1, pp. 54–66, 2012.

[13] L. Spinu, "Investigating the status of a rare cross-linguistic contrast: The case of Romanian palatalized postalveolars," *J. Acoust. Soc. Amer.*, vol. 143, no. 3, pp. 1235–1251, 2018.

[14] L. Spinu, A. Kochetov, and J. Lilley, "Acoustic classification of Russian plain and palatalized sibilant fricatives: Spectral vs. cepstral measures," *Speech Commun.*, vol. 100, pp. 41–45, 2010.

[15] P. Ghaffarvand Mokari and N. Mahdinezhad Sardhaei, "Predictive power of cepstral coefficients and spectral moments in the classification of Azerbaijani fricatives," *J. Acoust. Soc. Amer.*, vol. 147, no. 3, pp. EL228–EL234, 2020.

[16] J. Lilley, L. Spinu, and A. Athanasopoulou, "Exploring the front fricative contrast in Greek: A study of acoustic variability based on cepstral coefficients," *J. Int. Phonetic Assoc.*, vol. 51, no. 3, pp. 393–424, 2021.

[17] Y.-Y. Kong, A. Mullangi, and K. Kokkinakis, "Classification of fricative consonants for speech enhancement in hearing devices," *PLoS ONE*, vol. 9, no. 4, e95001, 2014. http://dx.doi.org/10.1371/journal.pone.0095001

[18] L. M. Jesus and P. J. Jackson, "Frication and voicing classification," in *8th Int. Conf. Comput. Process. Portuguese (PROPOR),* A. Teixeira, V. L. S. de Lima, L. C. de Oliveira, and P. Quaresma, Eds., Berlin, Heidelberg: Springer-Verlag, 2008, pp. 11–20.

[19] I. Vasilescu, B. Vieru, and L. Lamel, "Exploring pronunciation variants for Romanian speech-to-text transcription," in *SLTU 2014 – 4th Int. Workshop Spoken Lang. Technol. for Under-Resourced Lang.,* St. Petersburg, Russia, May 2014, pp. 161-168.

[20] I. Vasilescu, I. Chitoran, B. Vieru, M. Adda-Decker, M. Candea, L. Lamel, and O. Niculescu, "Studying variation in Romanian: deletion of the definite article -l in continuous speech," *Linguistic Vanguard*, vol. 5, no. 1, pp. 20170046, 2019.

[21] D. Pape and L. M. Jesus, "Devoicing of phonologically voiced obstruents: Is European Portuguese different from other Romance languages?" in *Proc. 17th Int. Congr. Phonetic Sci.*, Hong Kong, China, Aug. 2011, pp. 1566–1569.

[22] B. McMurray and A. Jongman, "What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations," *Psychol. Rev.*, vol. 118, pp. 219–246, 2011.

[23] *MATLAB and Statistics Toolbox.* (2013). The MathWorks, Inc.

[24] I. Chitoran, *The Phonology of Romanian: A Constraint-based Approach.* Berlin, Germany: Mouton de Gruyter, 2002, p. 292.

[25] J. Ohala, "The origin of sound patterns in vocal tract constraints," in *The Production of Speech,* P. F. MacNeilage, Ed. New York, NY: Springer, 1983, pp. 189–216.

[26] J. Ohala and C. Riordan, "Passive vocal tract enlargement during voiced stops," in *Rep. Phonol. Lab.,* no. 5, pp. 78–87, 1980.

[27] O. Fujimura, M. Macchi, and L. A. Streeter, "Perception of stop consonants with conflicting transitional cues: A cross-linguistic study," *Lang. and Speech,* vol. 21, pp. 337–346, 1978.

[28] J. J. Ohala, "The phonetics and phonology of aspects of assimilation," in *Between the Grammar and the Physics of Speech* (Papers in Laboratory Phonology I), J. Kingston and M. Beckman, Eds., Cambridge, MA, USA: Cambridge University Press, 1990, pp. 258–275.

[29] R. Wright, "A review of perceptual cues and cue robustness," in *Phonetically-Based Phonology,* B. Hayes, R. Kirchner, and D. Steriade, Eds., Cambridge, MA, USA: Cambridge University Press, 2004, pp. 34–57.

[30] D. Steriade, "Phonetics in phonology: The case of laryngeal neutralization," in *Papers in Phonology*, vol. 3 (UCLA Working Papers in Linguistics, no. 2), M. Gordon, Ed., Los Angeles: UCLA, 1999, pp. 25–146.

[31] N. L. Shin, "Women as leaders of language change: A qualification from the bilingual perspective," in *Proc. 6th Int. Workshop on Spanish Sociolinguistics*, Tucson, AZ, USA, Apr. 2013, pp. 135–147.

[32] W. Labov, (2001). *Principles of Linguistic Change, Vol. 2: Social Factors.* Oxford, U.K.: Blackwell, 2001, p. 592.