City University of New York (CUNY)

CUNY Academic Works

Publications and Research

Kingsborough Community College

2018

Acoustic Classification of Russian Plain and Palatalized Sibilant Fricatives: Spectral vs. Cepstral Measures

Laura Spinu

CUNY Kingsborough Community College

Alexei Kochetov

University of Toronto

Jason Lilley

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/164

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

**Acoustic classification of Russian plain and palatalized sibilant fricatives:
Spectral vs. cepstral measures**

Laura Spinu^{1*}, Alexei Kochetov¹ and Jason Lilley²

¹ Department of Communications & Performing Arts

Kingsborough Community College

City University of New York

Brooklyn, NY 11235

² Nemours Biomedical Research

Center for Pediatric and Auditory Speech Sciences

Wilmington, Delaware 19803

*e-mail: lspinu@kbcc.cuny.edu

Abstract

This study compares two methods for classifying voiceless sibilant fricatives forming a 4-way phonemic contrast found in Russian, but otherwise cross-linguistically rare. One method uses spectral measures, i.e. vowel formants, COG, duration and intensity of frication. The second method uses cepstral coefficients extracted from different regions inside fricatives and neighboring vowels. The corpus comprises 1,431 plain and palatalized fricatives from two places of articulation, produced by 10 speakers. Logistic regression was used to classify the productions of males and females together and separately. The productions of females yielded higher correct classification rates (highest 91.9%). Cepstral measures outperformed spectral measures across-the-board.

I. INTRODUCTION

The acoustic characteristics of fricatives have been the focus of numerous studies both before and after the early 2000s. While some studies sought to identify the general properties of different fricatives cross-linguistically (Gordon et al., 2002, Tabain and Watson, 1996), others narrowed down the focus of their study to a specific language (Jongman et al., 2000, Jesus and Shadle, 2002, 2003, Maniwa et al., 2009, among others), or even more specific aspects pertaining to fricatives within a language, such as whistled fricatives (Lee-Kim et al. 2014), laryngeal articulations (Nawrocki, 2008), pharyngeal articulation (Proctor et al., 2010), devoicing (Pape and Jesus, 2015), secondary palatalization (Spinu and Lilley, 2016, Kochetov 2017), or sibilance (Flipsen et al., 1999).

Russian is one of the few languages with a 4-way contrast involving palatalized sibilant fricatives, specifically: palatalized dental/alveolar /sʲ/, palatalized post-alveolar (prepalatal) /ʃʲ/, non-palatalized dental/alveolar /s/ and retroflex (apical post-alveolar) /ʂ/ (Timberlake, 2004). By investigating a contrast that is noteworthy in its rare cross-linguistic occurrence, the current study adds to the growing body of work on the acoustics of fricatives. Our goal is to identify the best methodological ways to classify fricatives accurately. We thus compare the performance of two classification methods, one based on spectral measures traditionally used in phonetic research and a novel method based on cepstral measures, on a corpus consisting of plain and palatalized Russian voiceless sibilant fricatives.

II. BACKGROUND

The most common acoustic measures previously used with fricatives include center of gravity, spectral peak location, spectral slope, spectral moments, noise duration, F2 onset frequency, static and dynamic amplitude measurements, and locus equations (see McMurray and Jongman, 2011, for a comprehensive review). While traditionally these measures were based on discrete Fourier transforms, multitaper spectra were recently introduced as better suited for stochastic parts of speech (Koenig et al., 2013, Lousada et al., 2012, Zygis et al., 2012). Much of the previous work on fricatives focused on identifying parameters that differ significantly between various categories (most commonly, place of articulation and/or voicing), but relatively few of them were designed specifically for classification, that is, in order to identify acoustic parameters able to reliably differentiate or discriminate a corpus of fricative consonants in terms of place of articulation and voicing.

Early studies performing classification of fricatives based on spectral moments yielded correct classification rates of 74.5%-77.7% (Forrest et al., 1988, with a set comprised of /f, θ, s, ʃ/) and 74%-78% (Tomiak, 1990, with /f, θ, s, ʃ, h/). Using spectral moments, duration, normalized amplitude, and spectral slope, Nissen et al. (2005) obtained a correct classification rate of 65% for adult productions of /f, θ, s, ʃ/. One of the most comprehensive recent classification studies employed a corpus of eight fricatives at four places of articulation, specifically /f, v, θ, ð, s, z, ʃ, ʒ/, produced by 20 English speakers

(McMurray and Jongman, 2011). Using 24 predictors comprising all the best known measures as well as newly-developed ones, the authors obtained correct classification rates between 79.2-85%, improving on Jongman et al.'s previous classification rate for the same corpus, i.e. 77% (Jongman et al., 2000).

Spinu and Lilley (2016) compared a novel method, based on cepstral coefficients, with a method based on spectral moments to classify 5 pairs of plain and palatalized Romanian fricatives (i.e. [f-f^j, v-v^j, z-z^j, ʃ-ʃ^j, x/h-ç^j]), produced by 31 native speakers and obtained a correct classification rate of 95.3%. Hidden Markov Models (HMMs) were used to divide each fricative into three regions based on their internal variance. The cues were extracted from regions inside the frication portion only, without any information from adjacent vowels. Crucially, their corpus did not include interdental fricatives, which have traditionally posed challenges to classification. The classification rate obtained may thus not be a major improvement over previous studies, but rather complementary to them, contributing data from a different part of the fricative 'landscape'. Even so, the high correct classification rates obtained recommend the cepstral coefficient/HMM-region method as a reliable way of identifying the crucial combination of a fricative's properties that makes it unique and distinguishable compared to all other fricatives. In the 2016 study, cepstral coefficients were shown to yield more accurate classification of place of articulation, palatalization, voicing and gender compared to spectral moments. No comparison was

made with more traditional fricative measurements, however, such as formants and center of gravity. The current study expands the comparison to these other measures.

III. CURRENT STUDY

In this study, we extend the method from Spinu and Lilley (2016) to a new language, Russian, focusing on a specific subset of fricatives – voiceless plain and palatalized sibilants. Adding to the complexity of the problem is the fact that this 4-way primary/secondary place contrast is extremely rare cross-linguistically. The absence or rarity of contrasts in languages is often associated with facts of phonetic difficulty and resulting phonological instability. This contrast therefore constitutes a good testing ground for the HMM-region cepstral-based classification method (HCCM).

Our corpus was originally collected for Kochetov (2017), where it was used to provide a general acoustic description of the Russian contrasts. The materials consisted of 48 target words with the fricatives /s/, /sʲ/, /ʂ/, and /ʃ/. The words were produced 3 times in a carrier phrase by 10 native speakers of Standard Russian (5 females and 5 males; median age 21.5), all born and raised in Russia, but at the time of the study residing in Canada. The resulting 1,431 tokens (144 tokens per speaker, minus 9 omissions) were annotated in Praat (Boersma and Weenink 2015), indicating the fricatives and preceding/following vowels. Please see Kochetov (2017) for additional details regarding the participants’ background, recording procedure, and annotation criteria.

A. Spectral measures

For the purposes of this study, a Praat script was run on the corpus data to obtain the following 16 spectral-based measurements from the annotated data:

Duration (in milliseconds) of the fricative and adjacent (word-internal) vowel

Amplitude (in dB) of the fricative and adjacent (word-internal) vowel

Centre of gravity of fricative noise (COG, or the first spectral moment, in Hz),

measured at 3 points in time: onset (C-on), midpoint (C-mid), and offset of the fricative (C-off), using a 25 ms Gaussian window and a 500 Hz to 10,000 Hz pass Hann filter. The windows were either aligned to fricative edges (C-on and C-off) or centred at the midpoint (C-mid). The low cutoff was set to exclude low-frequency room noise or voicing leakage from surrounding vowels, if any (cf. Zsiga, 2000; Nowak, 2006).

Formants F1, F2, and F3 (Hz) measured at 3 points within the following vowel (or the preceding vowel for word-final fricatives): onset (V-on), midpoint (V-mid), and offset (V-off), using a 25 ms Gaussian window and the Formant (Burg) algorithm. The windows were either aligned to fricative edges (V-on and V-off) or centred at the midpoint (V-mid).

B. Cepstral measures

The first 6 cepstral coefficients (c0-c5), Bark-scaled, were extracted from 10-ms frames inside each segment (fricative and adjacent vowels). HMMs were used to divide the segments into regions of internally minimized variance (Viterbi, 1967). Each HMM consists of three states arranged linearly. Each state models one region of a phoneme, and the state’s parameters comprise the means and variances of the feature vectors within that region. Each cepstral coefficient was averaged by region. Only the vocalic region adjacent to the fricative was used. The addition of vocalic data constitutes a notable difference compared to the previous analysis by Spinu and Lilley (2016).

C. Statistical analysis

For cepstral measures, 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 24 measures for each parameter set: 6 coefficients x 4 regions (3 consonantal regions + 1 vocalic region). Following McMurray and Jongman (2011), we conducted multinomial logistic regression analyses with consonant identity (s , s^j , ʃ , ʃ^j) as the dependent variable and the 24 measures as continuous explanatory variables. For spectral measures, consonant identity was used as the dependent variable and the 16 measures extracted as continuous explanatory variables. Logistic regression has been claimed to be best fitted for categorical response data, as theoretical problems arise with the application of discriminant analyses to them (Morrison and Kondaurova, 2009, p. 2160). We used Matlab R2013a (MATLAB, 2013) to determine

the regression coefficients. The first analyses were run on the entire data, including both genders. We then separated the corpus into male-only and female-only subcorpora and reran the analyses. We ran additional analyses on the Top 16 predictors from the cepstral set (in order to make it more comparable to the spectral set in terms of number of predictors), and the Top 5 predictors from both sets.

D. Results

Figure 1 shows the correct classification rates obtained with the cepstral and spectral set, using (1) all predictors, (2) the Top 16 predictors only, and (3) the Top 5 predictors only. These are shown for males and females combined, as well as for each gender separately. The set of the Top 5 predictors varied somewhat depending on the corpus used, specifically males, females, or males and females combined. For the latter corpus (both genders combined), the Top 5 spectral measures were F2-onset, COG-mid, C-intensity, COG-onset, and F3-onset. The Top 5 cepstral coefficients for the same set were C1.1.vowel, C4.3.fricative, C0.3.fricative, C2.1.vowel, C3.3.fricative (where C# stands for the coefficient, and the middle number stands for the region from which it was extracted). Note that for the spectral method, the 'all measures' set coincides with the 'Top 16' set, as a total of 16 measures had been extracted, hence the classification results are identical. It was only for the cepstral method that an advantage of using additional predictors was expected, which was borne out by the results. The highest classification rate (91.9%) was

yielded by the cepstral set using all 24 predictors and the female subcorpus. The lowest classification rate, specifically 63.9%, was obtained for male-only productions, with the spectral set using the Top 5 predictors only. Notably, the classification results across all measure sets improved between 8.2% and 13.9% when only the productions of females were included.

Table 1 provides the breakdown of correct classification rates by consonant for all analyses. Most error rates involve confusions between plain /ʃ/ and palatalized /ʃʲ/ (from 10.6% to 19.9%) and between plain /s/ and palatalized /sʲ/ (from 5.6% to 28.5%). That is, most errors involved the primary place of articulation. Errors across place of articulation (anterior versus posterior) were rare in the plain forms (0%-2.2%), but were encountered more frequently in the palatalized forms (from 0% to 12.8%), especially with the spectral method. Focusing on each gender, it was found that most of the classification errors encountered in the male set involved the plain-palatalized anterior pair /s/-/sʲ/, with error rates ranging from 11.4% to 28.5%. Secondary palatalization was a source of confusion for the posterior pair as well, but to a slightly lesser extent (14.4%-19.9%). These patterns are also present but somewhat less prominent in the female only results, particularly with respect to the anterior pair (5.6%-16.7%).

Figure 1: Correct classification rates by predictor set (all, Top 16, or Top 5) and method used (cepstral vs. spectral), for the productions of females (F), males (M) and both genders combined (MF).

IV. DISCUSSION

We set out to compare two methods for fricative classification and, in doing so, to provide descriptive data regarding the 4-way contrast of Russian voiceless sibilant fricatives. Our findings were that cepstral coefficients outperformed spectral measures, regardless of the measure type (spectral moments alone, as in Spinu and Lilley (2016), or other traditional fricative measures such as the ones explored here). Except for the classification of male productions, where the cepstral method outperformed the spectral

CC					SM				
MF	s	s ^j	ξ	∫ ^j	MF	s	s ^j	ξ	∫ ^j
s	82.50%	17.50%	0.00%	0.00%	s	73.00%	23.10%	1.90%	1.90%
s ^j	13.20%	85.10%	0.30%	1.40%	s ^j	11.20%	84.30%	1.10%	3.40%
ξ	0.80%	2.80%	79.50%	16.90%	ξ	1.70%	11.00%	70.20%	17.10%
∫ ^j	0.00%	4.70%	14.70%	80.60%	∫ ^j	3.10%	11.90%	16.40%	68.60%
M	s	s ^j	ξ	∫ ^j	M	s	s ^j	ξ	∫ ^j
s	82.70%	17.30%	0.00%	0.00%	s	69.30%	28.50%	1.70%	0.60%
s ^j	11.40%	88.10%	0.00%	0.60%	s ^j	15.90%	80.10%	0.60%	3.40%
ξ	0.60%	4.50%	77.30%	17.60%	ξ	1.70%	11.40%	67.00%	19.90%
∫ ^j	0.00%	5.60%	14.40%	80.00%	∫ ^j	2.20%	12.80%	16.10%	68.90%
F	s	s ^j	ξ	∫ ^j	F	s	s ^j	ξ	∫ ^j
s	93.30%	6.70%	0.00%	0.00%	s	80.00%	16.70%	2.20%	1.10%
s ^j	5.60%	94.40%	0.00%	0.00%	s ^j	9.40%	88.30%	0.00%	2.20%
ξ	0.00%	1.10%	84.40%	14.40%	ξ	0.60%	8.90%	78.90%	11.70%
∫ ^j	0.00%	1.10%	10.60%	88.30%	∫ ^j	1.70%	9.40%	13.90%	75.00%

Table 1: Classification table for Top 16 predictors, broken down by method (cepstral coefficients versus spectral measures) and gender (MF – both genders collapsed, M only, F only).

Note: CC stands for cepstral coefficients and SM for spectral measures.

method by 10.9%, the classifications yielded by the two sets were comparable when only the Top 5 predictors were used. An unexpected finding was that female productions tend to yield much higher correct classification rates.

With respect to the cepstral method outperforming the spectral measures, it should be pointed out that the potential advantages of this method have been discussed in the literature lately. Cepstral coefficients fared better than spectral moments in the classification of stop release bursts (Bunnell et al., 2004). In a study on the classification of voicing in fricatives in British English and European Portuguese (Jesus and Jackson, 2008), Mel-frequency cepstral coefficients (MFCC) were also used successfully. Furthermore, Ferragne and Pellegrino (2010) recommended MFCC 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). Kong et al. (2014) obtained 85% correct classification for three places of articulation in fricatives using a set of 13 MFCCs, suggesting that the advantages of cepstral coefficients carry over to segments other than vowels. Finally, Spinu and Lilley (2016) compared cepstral coefficients and spectral moments directly in a study classifying a corpus of 3,674 Romanian fricatives 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 regions inside a segment. While the advantage of using HMM-defined regions was small, it was consistently noted with most classifications attempted. The use of any type of regions, as opposed to averaging over an entire segment was shown to contribute substantially to the classification of palatalization. Thus, Spinu et al. (2012) obtained an overall classification rate of 78% based on the fricative corpus also used in Spinu and Lilley (2016). In the latter study, correct classification for palatalization was 88.2% with HMM regions and 87.3% with equal regions.

Dimension reduction is another consideration. When only the top three predictors were used, Spinu and Lilley (2016) found a larger drop in accuracy (compared to the use of all predictors) with the spectral moments (14%) than with the cepstral coefficients (10%). The difference between cepstral and spectral measures was not as large in the current study when the predictors were restricted to the Top 5, suggesting that the combination of spectral measures used here is better suited for the classification of fricatives and more robust than spectral moments alone. This may be due to the fact that palatalization in Russian is largely signaled by the acoustic information in the CV transitions (which was not included in Spinu and Lilley, 2016). A new classification with consonantal data alone reduced the overall correct classification rates based on spectral measures to 61.8% (from

74%), whereas there was a very slight increase (to 83.3% from 82.9%) for cepstral predictors. Both analyses were run on the MF corpus.

The gender differences we found may indicate a sociolinguistic asymmetry present in Russian, possibly connected to a mild tendency to neutralize the secondary palatalization contrast. This appears more widespread in (and perhaps even initiated by) men, and more prominent in anteriors. This tendency could reflect a certain degree of hypo-articulation on the part of the males, which may ultimately lead to the same result, specifically, sound change. These findings, however, should be further verified, preferably with monolingual Russian speakers, as it is not clear to what extent the subjects' English proficiency may have played a part - conceivably, one of the genders could be more fluent in English. The posterior plain-palatalized contrast too was prone to a high error rate, which may also indicate incipient neutralization. A much higher degree of confusability between Romanian posterior plain [j] and palatalized [jʲ] had been reported in Spinu and Liley (2016), where the plain form had been classified as palatalized over 50% of the time. In Russian, however, the plain form is realized as retroflex (apical post-alveolar [ʂ]), which likely makes it more acoustically dissimilar from its palatalized counterpart. The fact that robust and symmetrical differentiation of the Russian sibilant contrast was found provides further support for the acoustic/perceptual dispersion approach to fricative inventories (Padgett and Zygis, 2007, on Russian and Polish; Zygis and Padgett, 2010, on Polish).

V. CONCLUSIONS

Our study compared the performance of two sets of acoustic measures in the classification of Russian voiceless sibilants representative of a rare cross-linguistic contrast. Depending on the combination of measures and the specific subset of consonants used, classification was as high as 91.9% and as low as 63.9%. We have found that cepstral measures generally fare better than spectral ones in classification tasks, and also that the productions of females yield more successful classification rates across the board. Directions for further research include a study of Russian voiced fricatives or non-coronals, as well as a cross-linguistic study of fricatives using the same method.

Acknowledgements

To be added.

REFERENCES

- Baum, L.E., Petrie, T., Soules, G. and Weiss, N. **(1970)**. “A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains,” *Ann. Math. Statist.*, vol. 41, no. 1, 164-171.
- Boersma, P. and Weenink, D. **(2015)**. Praat: Doing phonetics by computer [Computer program]. Version 5.4.01 retrieved from <http://www.praat.org/>
- 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 Int. Conf. on Spoken Language Processing*. Jeju, Korea.

Ferragne, E., and Pellegrino, F. **(2010)**. "Vowel systems and accent similarity in the British Isles: Exploiting multidimensional acoustic distances in phonetics," *J. Phon.*, 38(4), 526-539.

Flipsen, P., Shriberg, L., Weismer, G., Karlsson, H., and McSweeney, J. **(1999)**. "Acoustic characteristics of /s/ in adolescents," *J. of Speech, Lang., and Hearing Research* 42, 663-677.

Forrest, K., Weismer, G., Milenkovic, P., and Dougall, R. **(1988)**. "Statistical analysis of word-initial voiceless obstruents: Preliminary data," *J. Int. Phon. Assoc.* 84, 115-124.

Gordon, M., Barthmaier, P. and Sands, K. **(2002)**. "A cross-linguistic acoustic study of voiceless fricatives," *J. Int. Phon. Assoc.* (2002), 32. Cambridge Univ. Press, 141-174.

Jesus, L.M. **(2001)**. "Acoustic Phonetics of European Portuguese Fricative Consonants," PhD thesis, School of Electronics and Comp. Science, Univ. of Southampton.

- Jesus, L. M. and Jackson, P. J. **(2008)**. “Frication and voicing classification,” In A. Teixeira, V. L. S. de Lima, L. C. de Oliveira, and P. Quaresma (Eds.), *Computational processing of the Portuguese language: The 8th Int. Conf. on Comp. Processing of Portuguese (PROPOR)* (pp.11-20). Berlin/Heidelberg: Springer Verlag.
- Jesus, L.M. and Shadle, C.H. **(2002)**. “A parametric study of the spectral characteristics of European Portuguese fricatives,” *J. Phon.* 30:3, 437-464.
- Jesus, L.M. and Shadle, C.H. **(2003)**. “Devoicing measures of European Portuguese fricatives,” In: Mamede, N., Baptista, J., Trancoso, I., Nunes, M. (eds.) *Comp. Processing of the Portuguese Language*, pp. 1-8. Springer, Heidelberg.
- Jongman, A., Wayland, R., and Wong, S. **(2000)**. “Acoustic characteristics of English fricatives,” *J. Acoust. Soc. America* 10/2000; 108 (3 Pt 1), 1252-1263.
- Kochetov, A. **(2017)**. “Acoustics of Russian voiceless sibilant fricatives,” *J. Int. Phon. Assoc.* 00/0, 1–28.
- Koenig, L.L., Shadle, C.H., Preston, J.L. and Mooshammer, C.R. **(2013)**. “Toward improved spectral measures of /s/: Results from adolescents,” *J. of Speech, Lang., and Hearing Research* 56: 1175-1189.
- Kong, Y.-Y., Mullangi, A., and Kokkinakis, K. **(2014)**. “Classification of fricative

consonants for speech enhancement in hearing devices," *PLoS ONE* 9(4): e95001.

doi:10.1371/journal.pone.0095001

Kuzla, C., Cho, T., and Ernestus, M. **(2007)**. "Prosodic strengthening of German fricatives in duration and assimilatory devoicing," *J. Phon.* 35 (3), 301-320.

Lee-Kim, S., Kawahara, S., and Lee, S. **(2014)**. "The 'whistled' fricative in Xitsonga: Its articulation and acoustics," *Phonetica*, 71, 50-81.

Lousada, M., L. Jesus, and D. Pape **(2012)**. "Estimation of stops' spectral place cues using multitaper techniques," *DELTA* 28(1), 1-26.

Maniwa, K., Jongman, A. and Wade, T. **(2009)**. "Acoustic characteristics of clearly spoken English fricatives," *J. Acoust. Soc. America* 125:6, 3962-3973.

MATLAB and Statistics Toolbox Release **(2013)**. Natick, MA: The MathWorks, Inc.

McMurray, B., and Jongman, A. **(2011)**. "What information is necessary for speech categorization? Harnessing variability in the speech signal by integrating cues computed relative to expectations," *Psych. Review* 118, 219-246.

Nawrocki, G. **(2008)**. "Laryngeal articulations of /x/ in Southern Polish," *ZAS Papers in Linguistics*, 49, 145 -165.

- Nissen, S. and Fox, R. **(2005)**. "Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective," *J. Acoust. Soc. America*, 118, 2570-2578.
- Nowak, P.M. **(2006)**. "The role of vowel transitions and frication noise in the perception of Polish sibilants," *J. Phon.*, 34 (2), 139-152.
- Padgett, J. and Zygis, M. **(2007)**. "The evolution of sibilants in Polish and Russian," *J. of Slavic Linguistics* 15(2), 291-327.
- Pape, D. and Jesus, L. **(2015)**. "Stop and fricative devoicing in European Portuguese, Italian and German," *Lang. and Speech*, 58(2), 224-246.
- Proctor, M. I., Shadle, C. H., and Iskarous, K. **(2010)**. "Pharyngeal articulation in the production of voiced and voiceless fricatives," *J. Acoust. Soc. America*, 127, 1507-1518.
- Shadle, C.H. and Mair, S.J. **(1996)**. "Quantifying spectral characteristics of fricatives," *Proceedings of the Int. Conf. on Spoken Language Processing (ICSLP 96)*, Philadelphia, USA, pp. 1517-1520.
- Spinu, L., Vogel, I. and Bunnell, T. **(2012)**. "Palatalization in Romanian – Acoustic properties and perception," *J. Phon.*, Vol. 40, Issue 1, 54-66.

- Spinu, L. and Lilley, J. **(2016)**. “A comparison of cepstral coefficients and spectral moments in the classification of Romanian fricatives,” *J. Phon.*, Vol. 57, 40-58.
- Sussman, H.M. **(1994)**. “The phonological reality of locus equations across manner class distinctions: Preliminary observations,” *Phonetica* 51, 119-131.
- Tabain, M. and Watson, C. **(1996)**. “Classification of fricatives,” *6th Australian Int. Conf. on Speech Science and Technology*. Adelaide: Australia, 623-628.
- Timberlake, A. **(2004)**. “*A reference grammar of Russian*,” Cambridge Univ. Press.
- Tomiak, G. **(1990)**. “*An acoustic and perceptual analysis of the spectral moments invariant with voiceless fricative obstruents*,” Doctoral Dissertation, SUNY Buffalo.
- Viterbi, A.J. **(1967)**. “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,” *IEEE Transactions on Inf. Theory* 13 (2): 260-269.
- Zsiga, E. C. **(2000)**. “Phonetic alignment constraints: Consonant overlap and palatalization in English and Russian,” *J. Phon.* 28(1), 69-102.
- Zygis, M. and Padgett, J. **(2010)**. “A perceptual study of Polish sibilants, and its implications for historical sound change,” *J. Phon.* 38(2), 207-226.
- Zygis, M., Pape, D., and Jesus, L. **(2012)**. “(Non)retroflex Slavic affricates and their motivation: Evidence from Czech and Polish,” *J. Int. Phon. Assoc.* 42(3), 281-329.