INTERSPEECH 2013

The phonological voicing contrast in Czech:
An EPG study of phonated and whispered fricatives

Radek Skarnitzl, Pavel Šturm, Pavel Machač

Institute of Phonetics, Faculty of Arts, Charles University in Prague, Czech Republic

`radek.skarnitzl@ff.cuni.cz, sturmp@seznam.cz, pavel.machac@ff.cuni.cz`

Abstract

Czech represents a language with a high correlation between phonological and phonetic voicing, but the perception of the voicing contrast is supported by other phonetic correlates. We employed electropalatography (EPG) to investigate the articulatory correlates of the voicing contrast in phonated and whispered speech. Eight subjects read short sentences and VCV sequences with alveolar/postalveolar fricatives as target sounds. The results revealed significant differences in the contact patterns of fortis (voiceless) and lenis (voiced) fricatives in phonated speech but the differences diminished in whisper. Implications of these findings are discussed with respect to general phonetic theory.

Index Terms: voicing, fortis/lenis, fricatives, whisper, EPG, Czech

1. Introduction

The voicing contrast distinguishes members of speech sound pairs like /p/-/b/ or /s/-/z/. While a glimpse at the IPA chart may suggest that the voicing contrast is a trivial matter of binary nature, corresponding to a mere "turning on and off" of the vocal folds, the phonetic reality is considerably more complex. First of all, it is necessary to distinguish between *phonological voicing* (i.e., specified in the language system) and *phonetic voicing* (i.e., the actual presence or absence of vocal fold vibration). Due to the lack of correspondence between the two in languages like English or German, where it is common for phonologically voiced obstruents to be partially or completely devoiced phonetically, the concept of *tenseness* was suggested to account for differences between sounds like /p, s/ (described as tense, fortis) on the one hand and /b, z/ (described as lax, lenis) on the other hand; for more detail see e.g. [1: 95]. However, tenseness turned out to be similarly problematic, since the predicted greater muscular effort in fortis obstruents was not supported by empirical evidence ([2], [3], [4]), and a search began for other correlates of the hypothesized tenseness in several domains. Thus correlates have been found in the temporal properties, with fortis obstruents being longer than lenis obstruents and pre-fortis vowels being shorter than pre-lenis vowels ([5], [6], [7]), in the spectral domain ([6], [8]), and in the area of articulation (see below); for a summary of the examinations of tenseness correlates see for instance [8], [9], [10]. It is the articulatory differences between fortis and lenis obstruents which constitute the focus of this paper.

The research into the articulatory correlates of the fortis/lenis contrast has generated results which were often inconsistent, statistically insignificant, or even contradictory – it is well known, in fact, that the variability of articulatory data is considerably higher than that of acoustic data. Thus, for example, the excursion and velocity of the articulators were found to be greater in the fortis [p] than in the lenis [b] [11] or in the fortis [k] than in the lenis [ɡ] [12], while a study by Ostry and colleagues [13] showed that the excursion and velocity was lower in the fortis [k] than in the lenis [ɡ]. A similar result was obtained using a more modern method, electromagnetic articulography (EMA) [14].

For the purpose of this study, we are particularly interested in how fortis and lenis consonants differ in their linguopalatal contact obtained by electropalatography (EPG). It is theoretically conceivable that a fortis obstruent like [t] will be produced with greater contact between the tongue and the palate than a lenis obstruent like [d]; that would, in turn, support the notion of articulation strength, of tenseness. Such a finding would also be in line with the strategy of enlarging the volume of supraglottal cavities documented in lenis plosives (see, e.g., [15], [16]), with the aim to facilitate the continuation of vocal fold vibration. However, it is equally conceivable that linguopalatal contact will be weaker in fortis obstruents, since the higher intraoral pressure will act on the surface of the tongue and "push it away" from the palate. The lower intraoral pressure in lenis obstruents would then mean that the tongue is laxer and can "spread around" more, resulting in a greater contact with the palate. Again, results supporting both these contradictory views have been published. On the one hand, Dixit compared the realizations of dental and retroflex Hindi plosives and found the linguopalatal contact to be greater in the fortis than in the lenis ones [17]. Similarly, Fuchs found a greater degree of contact in German [t] than [d] in two out of three analyzed speakers [18]. On the other hand, [19] revealed a greater contact in the lenis [d] than in the fortis [t] in English, and similar results were obtained for fricatives [20].

The aim of the present study is therefore to shed more light on the articulatory correlates of the phonological voicing contrast. We will examine linguopalatal patterns in Czech alveolar and postalveolar fricatives. Apart from analyzing normal, phonated speech, we will also focus on whispered speech. Research has shown that the difference between phonologically voiced and voiceless obstruents is preserved in whispered speech [21], [22], although it appears that especially the durational differences between fortis and lenis consonants (as mentioned above, lenis consonants tend to be shorter than fortis consonants) become blurred in whispered speech [22], [23]. Aerodynamic data also suggest that phonologically voiced whispered plosives lie "between" voiceless and (phonated) voiced plosives in terms of their intraoral pressure [24].

Czech belongs to the group of languages that manifest a strong correspondence between phonological and phonetic voicing, so that the presence of fundamental frequency (F_0) is the basic correlate of voicing in Czech. The results of [6]

25 – 29 August 2013, Lyon, France

nevertheless indicate that listeners are still able to provide above-chance performance in the identification of whispered fricatives in a VCV context. Thus, comparing the linguopalatal patterns of fortis and lenis consonants in both phonated and whispered speech might reveal interesting findings about the operation of the voicing mechanism in general.

2. Method

Linguopalatal contact was obtained, along with the acoustic signal, using the EPG3 system [25]; this system records the contact of the tongue with an individualized, custom-made artificial palate with the frequency of 100 Hz. EPG data were acquired from eight native speakers of Czech (four female, four male), mostly employees or students at the Institute of Phonetics in Prague (the first two authors of this study are included among the speakers). Prior to data acquisition, the speakers had been speaking with the artificial palate for at least 30 minutes, in line with generally acknowledged recommendations (see, e.g., [26]).

Due to the above-mentioned variability of articulatory data, the material of EPG research tends to be highly constrained; studies typically include minimal-pair pseudowords pronounced in isolation or embedded in carrier phrases. In this study, we analyzed the target consonants – /s/, /z/, /ʃ/ and /ʒ/ – in both these contexts. First, each fricative (F) was pronounced in the intervocalic position, VFV, where V included the short Czech monophthongs, /ɪ ɛ a o u/, and in short sentences in which the target fricative also appeared in the intervocalic position, at the beginning of a stressed syllable (e.g. *Ona s̲ama nejde*; *Kočka z̲ase mňouká*). Therefore, each speaker pronounced each fricative 60 times (5 vowels × 3 realizations × normal/phonated condition × VFV/sentences). In total, the linguopalatal contact was analyzed in 1,920 fricatives (8 speakers × 4 fricatives × 60 items per fricative).

In the ArticAssist software [25], the frame with the greatest degree of tongue-palate contact was identified for each fricative, and the number of contacts in the *target areas* for the two places of articulation was calculated. As expected, the obtained contact patterns displayed great variability but, importantly, they sometimes deviated markedly from the generally acknowledged EPG3 areas according to which the first two rows of the palate correspond to the alveolar place and the third and fourth rows to the postalveolar place [27]. Thus, the target areas had to be redefined to span the first three rows for the alveolars and from the second to the seventh row for the postalveolars, as illustrated in Figure 1.

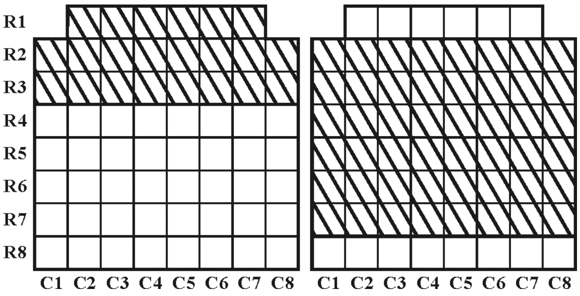

Figure 1: *Alveolar (left) and postalveolar (right) target areas (see text for more detail).*

The maximum number of electrodes that could be contacted was thus 22 for the alveolar [s] and [z] and 48 for the postalveolar [ʃ] and [ʒ]. In order to be able to compare fricatives across speakers, the absolute numbers of contacted electrodes in the target frames were further converted to normalized *z*-scores.

Several data reduction methods were proposed to facilitate statistical comparisons of linguopalatal contacts (see [27] for a review). Based on previous examinations of EPG patterns in fricatives [20], it appears that fortis and lenis fricatives may differ in the degree to which the contact extends into the central area of the artificial palate. That is why the centrality index [28] will also be used to compare the contact patterns, apart from the raw and normalized number of contacts. The formula for the centrality index (CC) is given by [28] as

$$CC = (\log\left(\left(1\left(\frac{C_1}{14}\right) + 17\left(\frac{C_2}{16}\right) + 289\left(\frac{C_3}{16}\right) + 4913\left(\frac{C_4}{16}\right)\right) + 1\right))/\log(5220 + 1)) \qquad (1)$$

where C_1 is the sum of electrodes contacted in columns 1 and 8, C_2 the sum contacted in columns 2 and 7, C_3 the sum in columns 3 and 6, and C_4 the sum contacted in columns 4 and 5. The index provides a measure of contact based on the differential involvement of central vs. lateral regions of the palate.

3. Results

In the first section we report the results concerning the degree of *linguopalatal contact* in the target areas (alveolar or postalveolar), as well as the *centrality index*. The following sections provide more detailed analyses of different experimental conditions and of individual speakers.

3.1. Voicing in phonated and whispered speech

The primary point of interest concerned the two contradictory possibilities regarding the degree of contact in fortis (voiceless) and lenis (voiced) fricatives, as mentioned in the Introduction. Figure 2 illustrates the normalized contact in the target areas (see Method). The results indicate a clear tendency towards greater contact in voiced fricatives, which is in agreement with the results of [19], [20].

Furthermore, the figure reveals that the differences between the phonologically voiceless and voiced cognates are slightly more pronounced for postalveolar [ʃ] and [ʒ] than for alveolar [s] and [z]. More importantly, the difference between the cognates only appears in phonated speech: similarly to the results reported in [22] and [23], the differences were diminished in whispered speech and the linguopalatal patterns were essentially identical in fortis and lenis fricatives, for both places of articulation. The interaction between VOICING (fortis *vs.* lenis) and MODE OF SPEECH (phonated *vs.* whispered) was highly significant for both alveolar ($F_{(1,\ 956)} = 10.00$, $p < 0.01$) and postalveolar ($F_{(1,\ 956)} = 10.78$, $p < 0.01$) fricatives. The statistical significance of the differences was computed through two multiple-factor ANOVAs and subsequent Tukey HSD post-hoc tests (marked in the figures with an asterisk).

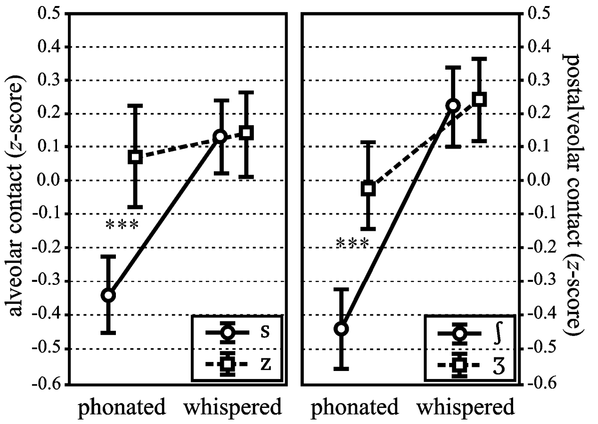

Figure 2: *Degree of contact (normalized z-score) in the target area of alveolar (left) and post-alveolar (right) fricatives in phonated and whispered speech. (*** $p < 0.001$ based on Tukey HSD post-hoc test.)*

The situation is analogous if we look at centrality. The phonologically voiced fricatives yielded a higher centrality index, i.e. they manifest a contact with the palate which extends more into the central regions of the palate than in the phonologically voiceless fricatives. As before, the voiced-voiceless differences were highly significant in phonated speech ($p < 0.01$), while in whisper they did not reach statistical significance. Similarly, postalveolar consonants show a more reliable difference than alveolar ones (interaction VOICING vs. MODE OF SPEECH: $F_{(1,\ 956)} = 4.22$, $p < 0.05$ for alveolar fricatives and $F_{(1,\ 956)} = 5.86$, $p < 0.05$ for postalveolar fricatives; differences in phonated speech as revealed by Tukey HSD: $p < 0.01$ and $p < 0.001$ for alveolar and post-alveolar fricatives, respectively).

3.2. Effect of material and vocalic context

Since articulation may differ markedly depending on the prosodic context, the target area contact was examined with respect to the TYPE OF MATERIAL in addition to VOICING and MODE OF SPEECH (ANOVA equations for all analyses are henceforward gathered in Table 1 at the end of the section). In phonated speech, the alveolar fricatives differed significantly (greater contact for [z] than [s]) both in sentences (Tukey HSD: $p < 0.01$) and in VCV sequences ($p < 0.05$). Predictably, the contact was significantly greater in the VCV context ($p < 0.001$), the more controlled material, which occurred irrespective of the voicing of the fricative. The postalveolar fricatives also formed greater contact in the voiced sound, but the difference was significant only for the VCV context ($p < 0.001$) since [ʃ] tended, surprisingly, towards greater contact in sentences than in the VCV material. As regards whispered speech, the VCV context involved greater contact than the sentence context in alveolar ($p < 0.01$) but not in postalveolar fricatives. Most importantly, in neither context did the fricatives differ with respect to phonological voicing.

As regards the type of material, it can be said that the centrality index behaves similarly to the overall contact patterns. With the exception of whispered [s], the centrality index was always greater for VCV than sentence contexts ($p < 0.05$ for phonated [s], $p < 0.01$ elsewhere). The voicing did not seem to have any marked influence on the centrality of the fricatives; the only significant difference was found between phonated [ʒ] and [ʃ] in the VCV context (the former having higher centrality index than the latter; $p < 0.001$).

In addition to the influence of the larger context, articulation of the fricatives was subject to variation according to the VOCALIC CONTEXT. The degree of contact tended to be greater with following close vowels and lesser with open vowels for both alveolar and postalveolar fricatives, and for both phonated and whispered speech. However, no differences between the fortis and the lenis consonants reached statistical significance.

3.3. Speaker variability

As is to be expected in articulatory research, the mean values displayed in Figure 2 hide a tremendous amount of individual variability. This is indicated in Figure 3 which compares the raw (i.e., not normalized) amount of alveolar and postalveolar contact for the individual speakers. The figure only shows results for phonated speech, but the pattern is nearly identical in whispered speech. It is interesting to note that while, for instance, Speaker 4 manifests by far the greatest degree of contact in alveolar fricatives, the degree of contact is about average in postalveolar fricatives, while the opposite applies for Speaker 7. In other words, the tendency towards greater or smaller linguopalatal contact generally holds across laryngeal conditions (phonated *vs.* whispered) in one speaker, but not across different places of articulation.

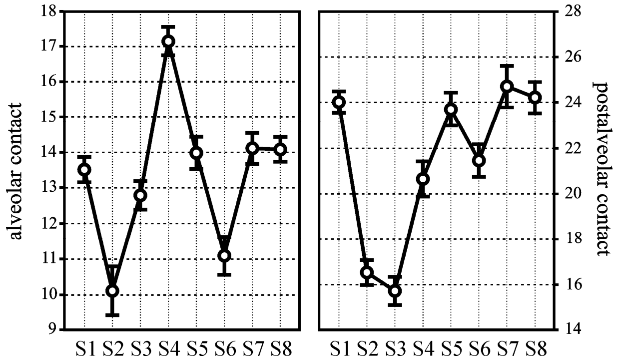

Figure 3: *Degree of contact (non-normalized number of contacted electrodes) in the target area of alveolar (left) and postalveolar (right) fricatives in phonated speech for the eight speakers.*

The analysis of phonological voicing contrasts in whispered speech did not reveal any significant differences between speakers (neither in the degree of contact nor in the centrality index). The contact patterns were nearly identical for the voiced/voiceless cognates in all speakers, with two non-significant exceptions (Speakers 3 and 4 in [ʃ]-[ʒ], where the phonologically voiced fricative was associated with greater contact); the only significant differences occurred, as ascertained above, between the speakers themselves due to the variability in their overall contact patterns.

However, the picture that emerges in phonated speech is more interesting. As can be seen from Figure 4, the centrality measure was able to significantly distinguish the alveolar fricatives [s z] in Speaker 8 (Tukey HSD: $p < 0.05$) and, with marginal significance, also in Speakers 5 and 6 ($p = 0.07$). To some extent, this seems to be reflected in the degree of contact as well (see right part of Figure 4), although the differences were not significant ($p > 0.1$): Speaker 6 was, along with Speakers 7 and 1, clearly distinguished, although the distinction was more blurred in Speakers 5 and 8. As regards the postalveolar place of articulation (not shown in the figure), one speaker (S3) had a markedly greater contact for voiced [ʒ] than for voiceless [ʃ] ($p < 0.01$), although the centrality index yielded only a marginally significant difference ($p < 0.1$). The centrality index also distinguished tenseness in Speaker 5 ($p < 0.1$) whose overall linguopalatal contact expressed by the z-score was, however, almost identical in the two consonants.

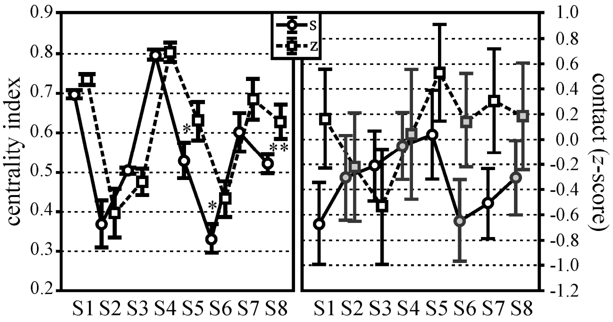

Figure 4: *Centrality index (left) and degree of contact (right) for individual speakers; data for the alveolar place of articulation in phonated speech. (* p < 0.10, ** p < 0.05 based on Tukey HSD post-hoc test.)*

place	**CC index**	**contact (*z*-score)**
/s z/	$F_{(1,\,952)} = 0.66, p = 0.42$	$F_{(1,\,952)} = 1.97, p = 0.16$
/ʃ ʒ/	$F_{(1,\,952)} = 1.93, p = 0.17$	$F_{(1,\,952)} = 2.56, p = 0.11$
/s z/	$F_{(4,\,940)} = 0.29, p = 0.88$	$F_{(4,\,940)} = 0.48, p = 0.75$
/ʃ ʒ/	$F_{(4,\,940)} = 0.32, p = 0.87$	$F_{(4,\,940)} = 1.05, p = 0.38$
/s z/	$F_{(7,\,928)} = 2.57, p < 0.05$	$F_{(7,\,928)} = 2.05, p < 0.05$
/ʃ ʒ/	$F_{(7,\,928)} = 1.93, p = 0.06$	$F_{(7,\,928)} = 0.94, p = 0.48$

Table 1: *ANOVAs for interactions between the effects of* VOICING *(fortis × lenis),* MODE OF SPEECH *(phonated × whispered) and one of the following:* TYPE OF MATERIAL *(upper part),* VOCALIC CONTEXT *(middle) and* SPEAKER *(lower part). Separately for alveolar and postalveolar place of articulation.*

4. Discussion and Conclusions

In this study, we analyzed the tongue-palate contact in voiceless-voiced (or fortis-lenis) pairs of fricatives in Czech. By focusing not only on phonated, but also on whispered speech, we are able to contribute with new findings to the discussion – which has been lively for several decades, as the sources cited in the Introduction show – concerning the correlates of phonological voicing.

Our results clearly suggest that linguopalatal contact is greater in the voiced (lenis) fricatives [z], [ʒ] than in their voiceless (fortis) counterparts [s], [ʃ], thereby supporting the theory according to which the mass of the tongue is displaced from the palate by the rising intraoral pressure in fortis sounds, while such an effect is considerably weaker in lenis sounds, allowing for greater tongue-palate contact (*cf.*, [19], [20]). The high correlation between the target area contact and the centrality index suggests that the midline groove is narrower in lenis than in fortis sounds.

What does this indicate about the correlates of tenseness, of the fortis-lenis distinction? It is interesting to realize that the arguments presented in literature in favour of stronger tongue-palate contact appear to have mentioned one half of the equation but not the other. Specifically, it is the more relaxed position of the tongue in lenis (indeed, lax) obstruents, resulting in the above-mentioned "spreading around" of the tongue mass. Put in slightly different words, lenis obstruents would appear to be associated with lower precision in attaining the articulatory target. The other half of the equation would then state that fortis (tense) obstruents will manifest greater articulatory precision, the natural consequence of which would be a lower degree of linguopalatal contact since the tongue muscles would be more sharply defined [Jan Volín, personal communication]. Thus, the differences in tongue-palate contacts between fortis and lenis obstruents may be regarded as articulatory correlates of tenseness. In a yet more complex explanation, however, plosives and fricatives may behave differently in this respect; that might be related to the fact that plosives have been demonstrated to have a virtual articulatory target [29], while it is clear that we attain the target during the production of fricatives.

However, the highly significant difference is only maintained in phonated speech (see Figure 2), whereas in whispered speech it is essentially lost. We may hypothesize, therefore, that the difference observed in phonated speech is associated with the actual presence or absence of (phonetic) voicing (i.e., presence or absence of vocal fold vibration) and the aerodynamic conditions which are related to it.

The results of our analyses also confirm the high variability of articulatory data. Moreover, it is obvious from the comparison of alveolar and postalveolar contact that speakers acquire idiosyncratic articulatory strategies for the production of individual speech sounds. Although results of an articulatory study are not directly applicable in forensic contexts, it would be interesting to compare the tendencies observed in the EPG patterns for individual speakers with the individual patterning of acoustic data, specifically spectral parameters. Such a detailed analysis of articulatory patterns may lead to the revelation of fine differences on the acoustic level.

To conclude, our EPG data from eight speakers confirm the differential articulation of fortis (voiceless) and lenis (voiced) sounds for phonated Czech. However, direct comparison of EPG with acoustic and perceptual data is needed in future experiments for a more complete picture of the phonological voicing contrast in both phonated and whispered speech.

5. Acknowledgements

The research was supported by projects GAČR 406/12/0300 and 406/12/0298. The authors would also like to thank the anonymous reviewers for their helpful comments.

6. References

[1] Ladefoged, P. and Maddieson, I., "Sounds of the World's Languages", Oxford: Blackwell, 1996.

[2] Ladefoged, P., "Preliminaries to Linguistic Phonetics", Chicago: University of Chicago Press, 1971.

[3] Malécot, A., "Mechanical pressure as an index of 'force of articulation'", Phonetica, 14: 169-180, 1966.

[4] Malécot, A., "The lenis-fortis opposition: its physiological correlates", J Acoust Soc Am, 47: 1588-1592, 1969.

[5] Lisker, L., "Closure duration and the intervocalic voiced-voiceless distinction in English", Language, 33: 42-49, 1957.

[6] Slis, I. H. and Cohen, A., "On the complex regulating the voiced-voiceless distinction I", Language and Speech, 12: 80-102, 1969.

[7] Chen, M., "Vowel length variation as a function of the voicing of the consonant environment", Phonetica, 22: 129-159, 1970.

[8] Jessen, M., "Phonetics and phonology of tense and lax obstruents in German", Amsterdam: John Benjamins, 1998.

[9] Kohler, K. J., "Phonetic explanation in phonology: the feature fortis/lenis", Phonetica, 41: 150-174, 1984.

[10] Skarnitzl, R., "The voicing contrast not only in Czech", Praha: Epocha, 2011. [in Czech]

[11] Sussman, H. M., MacNeilage, P. F. and Hanson, R. J., "Labial and mandibular dynamics during the production of bilabial consonants: Preliminary observations", J Speech and Hearing Res, 16: 397-420, 1973.

[12] Kohler, K. J., "Timing of articulatory control in the production of plosives", Phonetica, 38: 116-125, 1980.

[13] Ostry, D. J., Keller, E. and Parush, A., "Similarities in the control of the speech articulators and the limbs: Kinematics of tongue dorsum movement in speech", J Exp Psych: Hum Perc and Perf, 9: 622-636, 1983.

[14] Löfqvist, A. and Gracco, V. L., "Tongue body kinematics in velar stop production: Influences of consonant voicing and vowel context", Phonetica, 51: 52-67, 1994.

[15] Westbury, J. H., "Enlargement of the supraglottal cavity and its relation to stop consonant voicing", J Acoust Soc Am, 73: 1322-1336, 1983.

[16] Svirsky, M. A., Stevens, K. N., Matthies, M. L., Manzella, J., Perkell, J. S. and Wilhelms-Tricarico, R., "Tongue surface displacement during bilabial stops", J Acoust Soc Am, 102: 562-571, 1997.

[17] Dixit, R. P., "Linguotectal contact patterns in the dental and retroflex stops of Hindi", J of Phon, 18: 189-201, 1990.

[18] Fuchs, S., "Articulatory correlates of the voicing contrast in alveolar obstruent production in German", Berlin: ZAS Papers in Linguistics 41, 2005.

[19] Dagenais, P. A., Lorendo, L. C. and McCutcheon, M. J., "A study of voicing and context effects upon consonant linguopalatal contact patterns", J of Phon, 22: 225-238, 1994.

[20] McLeod, S., Roberts, A. and Sita, J., "Tongue/palate contact for the production of /s/ and /z/", Clin Ling Phon, 20: 51-66, 2006.

[21] Ito, T., Takeda, K. and Itakura, F., "Analysis and recognition of whispered speech", Speech Comm, 45: 139-152, 2005.

[22] Jovičić, S. T. and Šarić, Z., "Acoustic analysis of consonants in whispered speech", J Voice, 22: 263-274, 2008.

[23] Machač, P. and Šturm, P., "The phonological contrast of voicing in whispered Czech and its phonetic correlates – A preliminary study", in R. Vích [Ed], 20th Czech-German Workshop - Speech Processing, 34-43, 2010.

[24] Higashikawa, M., "Perceptual, acoustical and aerodynamic study of whispering", Nippon Jibiinkoka Gakkai Kaiho, 97: 1268-80, 1994. [abstract]

[25] Wrench, A., "Artic Assist (version 1.14)", Edinburgh: Queen Margaret University College, 2006.

[26] Hardcastle, W. J. and Gibbon, F., "Electropalatography and its clinical applications", in M. J. Ball and C. Code [Eds], Instrumental clinical phonetics, 149-193, London: Whurr Publishers, 1997.

[27] Gibbon, F. and Nicolaidis, K., "Palatography", in W. J. Hardcastle and N. Hewlett [Eds], Coarticulation: Theory, Data, and Techniques, 229-245. Cambridge: Cambridge University Press, 1999.

[28] Fontdevila, J., Pallarès, M. D. and Recasens, D., "The contact index method of electropalatographic data reduction", J Phon, 22: 141-154, 1994.

[29] Löfqvist, A. and Gracco, V. L., "Lip and jaw kinematics in bilabial stop consonant production", Journal of Speech, Language and Hearing Research, 40: 877-893, 1997.