5. SIMPLE PHONOLOGICAL CONTRASTS CAUSING COMPLEX PHONETIC DIFFERENCES

Phil HOWSON – Philip J. MONAHAN

5.1. Introduction

Feature-based phonological theories posit distinctive features as the fundamental representational unit (Chomsky & Halle, 1968; Clements, 1985). These features describe major segmental classes (e.g., sonorant, approximant), states of the larynx (e.g., voice, spread glottis), manner (e.g., nasal, strident) and place of articulation (e.g., coronal, dorsal). Feature labels are based on articulatory- and acoustic-phonetic characteristics of the segments they describe and are assigned a value, where "+" denotes the presence and "–" the absence of the feature. For example, a segment is [+nasal] if the velum lowers during articulation allowing airflow to pass through the nasal cavity. While distinctive features are powerful in accounting for phonological patterns, the phonetic realization of speech sound segments is more complex than the features assigned to them.

Complex interactions of the speech articulators are often reduced to a simple ± feature contrast within the phonology. A prime example of this complexity is found in voicing opposition, represented as [±voice], which reflects either the presence or absence of vocal fold vibrations; however, the voicing contrast is one of the most commonly investigated phonetic issues (Fuchs, 2005), and multiple acoustic cues exist to voicing, rather than simply the presence or absence of vocal fold vibration (Liker & Gibbon, 2013). These cues have been shown to reflect the activity of a number of physiological mechanisms that result in voicing being a complex issue.

To produce voicing, a transglottal pressure differential is necessary, such that there is greater subglottal than supraglottal pressure (Ohala & Solé, 2010). This difference creates vibration of the vocal folds as the air flows from the lungs into the oral cavity. Creating this difference is largely the product of manipulating the shape and size of the supraglottal cavities (Liker & Gibbon, 2013). Achieving the difference in supraglottal shape and size leads to complex differences in the articulation of voiced and voiceless pairs.

Differences in articulation have been successfully studied using a variety of research techniques: electropalatography (EPG; Dixit & Hoffman, 2004; Recasens & Espinosa, 2007; Liker & Gibbon, 2011; see also Chapter 3 in this book), magnetic resonance imaging (MRI; Proctor, Shadle & Iskarous, 2010), electromagnetic articulography (EMA; Fuchs, 2005), and acoustic measures (Coleman, 2003). This work has revealed complex activities with respect to a number of articulatory specifications. During the production of voiced fricatives, for example, EPG studies reveal more anterior contact and smaller groove width, compared to voiceless fricatives (Dagenais, Lorendo & McCutcheon, 1994; Dixit & Hoffman, 2004). Liker & Gibbon (2013) also report that tongue-palate contact in voiceless fricatives is delayed significantly in the achievement of anterior contact compared to voiced fricatives. Dorsal contact, however, was not delayed for voiceless fricatives. The authors concluded that the closure was closer to the rear pharynx and created an air-pressure control mechanism, which is required during the production of voiceless fricatives. Similarly, Proctor et al. (2010) found that the production of voiced fricatives resulted in a larger pharyngeal volume using MRI, which they suggest mediates the transglottal pressure necessary to facilitate voicing. The largest part of this displacement was found with tongue advancement. That is to say that the tongue dorsum for voiceless fricatives involved a significant amount of retraction, which resulted in a greatly reduced pharyngeal volume compared to the voiced fricatives. Proctor et al. (2010) suggested that this tongue retraction for voiceless fricatives is an active air-pressure control mechanism and reduces the increased airflow due to the open glottis. Taken together, this work reveals complex articulatory interactions mediate the binary phonological distinction.

The purpose of this study is to examine the voicing contrast in Czech fricatives using acoustic measures. Czech is a particularly good candidate for the study of fricative voicing given its elaborate fricative inventory. Fine grained differences in the formant transitions and spectral measures will be compared to test the hypothesis that there is a retracted tongue dorsum during the production of voiceless fricatives to maintain air-pressure control (Proctor et al., 2010).

5.2. Method

5.2.1. Participants

Four native speakers of Czech (two male [C1, C2]; two female [C3, C4]; mean age = 22 years old) were recruited from the Toronto area. All speakers had been raised in the Czech Republic until at least the age of 18, had no self-reported speaking or hearing impairments and provided written informed consent prior to participation.

5.2.2. Materials

Words were chosen such that the relevant fricative of interest was located in word-initial position followed by the low vowel /a/. The data set included the following real Czech words: *sát* [saːt] suck, *záď* [zaːc] stern (ship), *řád* [r̝aːt] order, *šál* [ʃaːl] scarf, *žák* [ʒaːk] pupil (school). The fricatives of interest were the voiceless [s] and [ʃ] and the voiced [z], [ʒ] and [r̝].

5.2.3. Procedure

Recordings were collected in a sound attenuated booth at the University of Toronto using a Lavaliere microphone, with a sampling frequency of 48 kHz and bit-depth of 32-bits. Participants first read "The North Wind and the Sun" (Dankovičová, 1999) in Czech twice to habituate to speaking in Czech. Single words were then presented in a randomized order on a computer screen. Each target was presented with a number of filler tokens. There were a total of 21 filler tokens each with the same environment but with different consonants. As such, the list contained all consonant contrasts in Czech. Each token was read aloud three times. The list was produced twice for a total of six tokens of each phoneme for each participant (24 total tokens for each phoneme, 120 total tokens).

5.2.4. Data Analysis

Data measurements were taken at two specific locations. Formant measures (F1 and F2) were taken at 5% into the vowel following each fricative, using a Burg analysis in Praat (v. 5.4.01; Boersma & Weenink, 2014). The onset of the vowel was taken to be the absence of frication accompanied by a clearly visible glottal pulse in the spectrogram. F2–F1 (Sproat & Fujimura, 1993) measures were obtained by subtracting F1 from F2. Centre of gravity (COG) measures were taken at the mid-point of the fricative, using a 10 ms window in Praat. Spectral peaks were calculated from a Thomson multi-taper spectral density estimate (Thomson, 1982) using

MATLAB (2009a). The peak measure for each voiced and voiceless token was measured (Figure 5–1). Spectral measures below 1 kHz were ignored to avoid glottal energy skewing the measurements in the voiced tokens. Frequencies above 12 kHz were ignored because energy generated from the front cavity is found within this range (Żygis, Pape & Jesus, 2012).

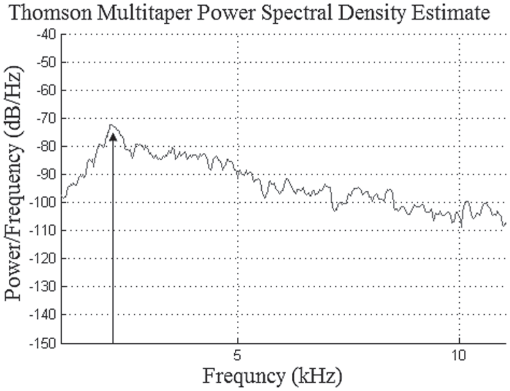

Figure 5–1: Thomson multi-taper spectral density estimate for a token of /ʃ/ produced by C1. The arrow indicates the location of the spectral peak.

5.3. Results

Inferential statistics were performed with linear mixed effects models using the function lme() in the R statistical environment (R Development Core Team, 2013). This method was adopted due to the unbalanced nature of the design, i.e., more Voiced than Voiceless observations. The fixed effects structure included the two-level factor Voicing (Voiced, Voiceless) and the random effects structure included random by-subject slopes for Voicing. Post-hoc analyses to compare differences between individual fricative categories were calculated using generalized linear hypothesis testing using the function glht() with a Tukey-correction for multiple comparisons. Spectral values were log-transformed prior to statistical analysis. First, the results for the formant measures are presented, followed by the spectral measures (COG and spectral peak).

Main effects of Voicing were observed for F1 [Voiced μ: 522 Hz; Voiceless μ: 600 Hz; $\beta = -0.15$, SE = 0.02; t = –6.39; $p < 0.01$], F2 [Voiced μ: 1636 Hz; Voiceless μ: 1526 Hz; $\beta = 0.07$, SE = 0.02; t = 3.67; $p < 0.01$] and the difference between F2 and F1 [Voiced μ_{Δ}: 1117; Voiceless μ_{Δ}: 926; $\beta = 0.19$, SE = 0.04; t = 5.23; $p < 0.01$].

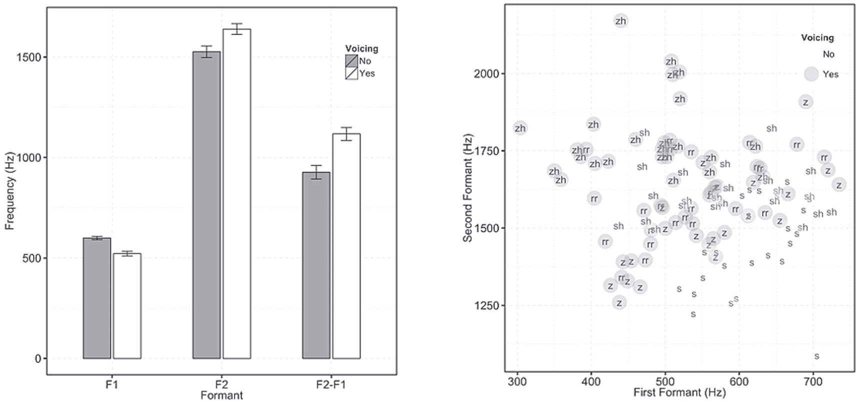

Figure 5–2: (A) Comparison of the first (F1) and second (F2) formants, as well as the difference between the two (F2–F1) by voicing-type. Error bars represent the standard error of the mean. (B) Scatter plot the distribution of fricatives as a function of the following F1 an F2 formant values of the vowel. Tokens in grey circles represented voiced sounds. (“s” = [s]; “sh” = [ʃ]; “z” = [z]; “zh” = [ʒ]; “rr” = [r̥])

That is, F1 was reliably higher (at its 5% point) immediately following voiceless fricatives, while F2 was higher and the F2 minus F1 difference was larger immediately following voiced fricatives (see Figure 5–2). Visualizing the distribution of fricatives by their F1 and F2 measures highlights the observations above: Voiceless fricatives co-occurred with a lower F1 and higher F2. Paired comparisons for each of the fricatives were also performed. These are provided in Table 5–1.

Pair	F1			F2			F2-F1		
	β	SE	z	β	SE	z	β	SE	z
[s]–[r̥]	0.16	0.03	4.89***	-0.11	0.02	-6.41***	-0.28	0.04	-7.91***
[ʃ]–[r̥]	0.09	0.03	3.09*	0.01	0.03	0.32	-0.03	0.06	-0.55
[z]–[r̥]	0.05	0.03	1.63	-0.05	0.03	-2.03	-0.11	0.04	-2.48
[ʒ]–[r̥]	-0.12	0.04	-3.22*	0.11	0.03	3.36**	0.21	0.05	3.92***
[ʃ]–[s]	-0.07	0.04	-1.66	0.12	0.04	3.31**	0.24	0.07	3.55**
[z]–[s]	-0.10	0.04	-2.50	0.06	0.03	2.06	0.17	0.05	3.42**
[ʒ]–[s]	-0.28	0.04	-6.57***	0.23	0.04	6.10***	0.49	0.06	7.55***
[z]–[ʃ]	-0.04	0.03	-1.50	-0.06	0.05	-1.35	-0.07	0.07	-0.99
[ʒ]–[ʃ]	-0.21	0.04	-5.58***	0.10	0.02	5.49***	0.24	0.05	5.27***
[ʒ]–[z]	-0.17	0.04	-4.58***	0.17	0.04	3.85**	0.31	0.05	5.99***

Table 5–1: Model output of the pairwise comparisons for F1, F2 and the difference between the two. *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.\ 05$.

The post-alveolar [ʒ] is reliably different from all other categories across the three spectral vowel measurements. In the scatter plot in Figure 5–2, it is evident that [ʒ] has the lowest F1 and highest F2 values of the other categories.

When analyzing the spectral center of gravity, we find no reliable main effect of voicing [Voiced μ: 5501 Hz; Voiceless μ: 5892 Hz; β = −0.07, SE = 0.06; t = −1.19; $p = 0.24$]. There was, however, a significant difference in voicing for spectral peak [Voiced μ: −74.9 dB; Voiceless μ: −71.2dB; β = −3.72, SE = 1.02; t = −3.65; $p < 0.01$][1] (Figure 5–3).

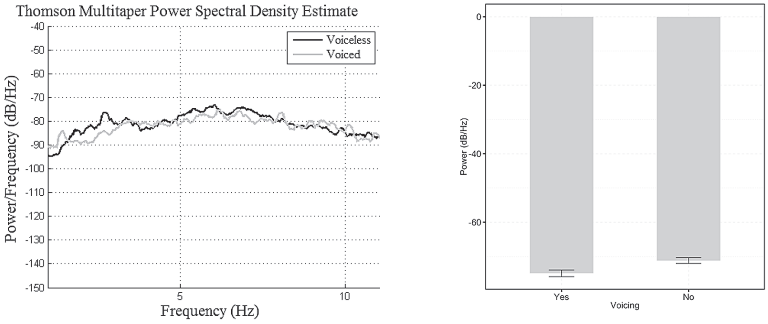

Figure 5–3: (A) Thomson multitaper power spectral density estimate averages for voiceless and voiced token. (B) Results of the spectral peak analysis. Error bars represent the standard error of the mean.

Table 5–2 presents the pairwise comparisons for the spectral peak.

Pair	F1		
	β	SE	z
[s]–[ɾ̞]	-1.97	2.29	-0.86
[ʃ]–[ɾ̞]	-0.19	1.54	-0.12
[z]–[ɾ̞]	-8.05	2.61	-3.09*
[ʒ]–[ɾ̞]	-6.36	1.58	-4.02***
[ʃ]–[s]	1.78	1.58	1.13
[z]–[s]	-6.08	1.39	-4.39***
[ʒ]–[s]	-4.40	1.65	-2.67
[z]–[ʃ]	-7.86	1.90	-4.15***
[ʒ]–[ʃ]	-6.18	1.29	-4.79***
[ʒ]–[z]	1.68	1.77	0.95

Table 5–2: Pairwise comparisons for spectral peak (dB/Hz). *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$.

---

[1] Given that the spectral peak values are all negative, statistical analysis was performed on the raw, non-log transformed values.

5.4. Discussion

The results demonstrate differences in the articulation of voiceless and voiced fricatives in Czech. They differ in F1, F2 and F2–F1 transitions, as well as spectral peaks. The F1 results show that voiceless consonants are articulated lower than voiced consonants are. The raised tongue observed in voiced consonants mirrors EPG findings, which show more lingual tongue-palate contact for voiced fricatives (Dixit & Hoffman, 2004; Recasens & Espinosa, 2007; Liker & Gibbon, 2011; Skarnitzl, Šturm & Machač, 2013). Overall, the raised tongue suggests a smaller pre-constriction vocal tract volume, which helps achieve the necessary transglottal pressure differential necessary to facilitate both voicing and frication. The F2 and F2–F1 both indicate that voiceless fricatives in Czech are articulated further back in the mouth than voiced consonants. This suggests a reduced pharyngeal volume for the articulation of voiceless fricatives in Czech. This result is expected if a retracted tongue dorsum is required to mediate the air pressure during the articulation of voiceless fricatives (Proctor et al., 2010). A specific advantage of the finding of this study is that the participants were not in a supine position, which is required for MRI studies. This indicates without a doubt that the retracted dorsum is not a result of gravity.

Interestingly, the increased airflow, which results from an open glottis (Ohala & Solé, 2010), modulates the intensity of the frication. This is manifested in the current findings in that the spectral peak for voiceless consonants was notably higher, suggesting more intensity produced during frication as a result of the increased airflow. This is likely the result of the airflow hitting the teeth at an increased velocity.

Another question raised is the extent to which phonological systems should account for patterns like those observed in our findings. Feature based theories are descriptively powerful and can account for a wide variety of phonological phenomena. In our results, however, we observed acoustic properties that divulge complex articulatory interactions. Knowledge of such small phonetic differences is not typically encoded in the features, yet these interactions appear to be necessary, raising the issue of how such knowledge should be representationally encoded.

5.5. Conclusion

This chapter used acoustic measures to show differences in formant transitions and spectral peaks for voiceless and voiced fricatives in Czech. The results indicated that voiceless fricatives are articulated lower and further back than the voiced counterparts. Voiceless fricatives also had a higher spectral peak than their voiced counter parts. The results support previous findings that the voiced fricatives have a smaller pre-constriction volume that helps achieve the pressure differential required for both voicing and frication. Furthermore, the results suggest the retracted tongue observed for voiceless fricatives is to mediate the airflow produced from the open glottis. The difference in airflow, however, is not completely nullified as the increased airflow still results in an increase in intensity (a higher spectral peak).

References

Boersma, P. – Weenink, D. (2014), *Praat: doing phonetics by computer* [Computer program], Version 5. 4. 01, http://www.praat.org/.

Chomsky, N. – Halle, M. (1968), *The sound pattern of English*, New York: Harper and Row.

Clements, G. N. (1985), The geometry of phonological features, *Phonology* 2(1), pp. 225–252.

Coleman, J. (2013), Discovering the acoustic correlates of phonological contrasts, *Journal of Phonetics* 31(3–4), pp. 351–372.

Dagenais, P. – Lorendo, L. – McCutcheon, M. (1994), A study of voicing and contexteffects upon consonant linguapalatal contact patterns, *Journal of Phoentics* 22, pp. 225–238.

Dankovičová, J. (1999), Czech, *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, pp. 70–74, Cambridge: Cambridge University Press.

Dixit, R. – Hoffman, P. (2004), Articulatory characteristics of fricatives and affricates in Hindi: An electropalatographic study, *Journal of the International Phonetic Association* 34(2), pp. 141–160.

Fuchs, S. (2005), Articulatory correlates of the voicing contrast in alveolar obstruent production in German, *ZAS Working Papers in Linguistics* 41.

Liker, M. – Gibbon, F. (2011), Groove width in Croatian voiced and voiceless postalveolar fricatives. In Lee, W. – Zee, E. (eds.), *The 17th International Congress of Phonetic Sciences* (ICPhS17), pp. 1238–1241, Hong Kong: University of Hong Kong, China.

Liker, M. – Gibbon, F. (2013), Differences in EPG contact dynamics between voiced and voiceless lingual fricatives, *Journal of the International Phonetic Association* 43(1), pp. 49–64.

MATLAB 9.0 and Statistics Toolbox 9.1, The MathWorks, Inc. Natick, Massachusetts, United States.

Ohala, J. – Solé, M. (2010), Turbulence and phonology. In Fuchs, S. – Toda, M. – Żygis, M. (eds.), *Turbulent sounds: An interdisciplinary guide*, pp. 37–102, Berlin & New York: De Gruyter Mouton.

Proctor, M. – Shadle, C. – Iskarous, K. (2010), Pharyngeal articulation in the production of voiced and voiceless fricatives, *Journal of the Acoustical Society of America* 127(3), pp. 1507–1518.

R Development Core Team (2013), *R: A language and environment for statistical computing*, http://www.R-project.org/.

Recasens, D. – Espinosa, A. (2007), An electropalatographic and acoustic study of affricates and fricatives in two Catalan dialects, *Journal of the International Phonetic Association* 37(2), pp. 143–172.

Skarnitzl, R. – Šturm, P. – Machač, P. (2013), The phonological voicing contrast in Czech: An EPG study of phonated and whispered fricatives, *Proceedings of Interspeech 2013*, pp. 3191–3195.

Sproat, R. – Fujimura, O. (1993), Allophonic variation in English /l/ and its implications for phonetic implementation, *Journal of Phonetics* 21, pp. 291–311.

Żygis, M. – Pape, D. – Jesus, L. M. T. (2012), (Non-)retroflex Slavic affricates and their motivation: Evidence from Czech and Polish, *Journal of the International Phonetic Association* 42, pp. 281–329.