An acoustic examination of the three-way sibilant contrast in Lower Sorbian

Phil Howson

1 The University of Toronto, Canada
phil.howson@mail.utoronto.ca

Abstract

The current paper presents an acoustic study of Lower Sorbian. Four native speakers of Lower Sorbian participated in this study. Center of gravity, standard deviation, skewness, kurtosis and slope measures were taken to measure the fricatives' spectral qualities. Vocalic transitions were also taken to gather finer grained information about the sibilants' articulation. The results indicated there is a three-way contrast in Lower Sorbian. The alveolar and retroflex segments have approximately the same place of articulation, but the overall tongue shape contributes significantly to the difference in spectral properties. The alveolopalatal segment is composed of a tongue-palate contact which extends from the dental place towards the palatal. This suggests that the length of the constriction also greatly affects the spectral properties of the fricatives. However, the phonetic realization of these segments is not the same as in Polish. The COG measures indicate both the Lower Sorbian retroflex and alveolopalatal segments is much lower than Polish. This suggests that the phonetic instantiation of the three-way contrast is variable. The results also indicate that the acoustic-perceptual cues from the vowel transitions and fricative spectra work in conjunction to distinguish the sibilants because no single cue is reliable enough to distinguish all of them.

Index Terms: Lower Sorbian, acoustics, three-way sibilant contrast, articulation

1. Introduction

Lower Sorbian has a cross-linguistically uncommon three-way contrast among the sibilant fricatives. It contrasts at the alveolar /s, z/, retroflex /ʃ, z/ and alveolopalatal /ɕ, z/ place of articulation [1]. Žygis [1] argues that the acoustic-perceptual relationships determine the sibilant inventory of languages. Her cross-linguistic examination Polish, Bulgarian and Russian revealed that the acoustic distance of /ʃ/ from /s/ is sufficient to distinguish the two in two-way sibilant systems; however, /ʃ/ has a relatively small acoustic distance from both /s/ and /z/ when considered in conjunction. Žygis [1] posits that /ʃ/ is not optimal for maintaining a three-way contrast and phonological systems are much more likely to employ a retroflex and alveolopalatal contrast to maintain as much perceptual distance as possible. Žygis, Pape, & Jesus [2] performed an acoustic study of the Czech post-alveolar affricate, /ť/, and the Polish alveolopalatal affricate, /ś/ and came to the same conclusion based on acoustic data. Žygis et al. [2] state that the acoustic cues for each of the affricates are too similar and that there is perceptual motivation behind sound changes from post-alveolar to retroflex in languages with three-way contrasts. This conclusion is contra Hall [3] who suggests that the reason post-alveolar and alveolopalatal fricatives do not co-occur in the same inventory is because the feature specification for both phonemes is identical: [+coronal, -anterior, +distributed].

Nowak [4] performed an acoustic analysis and perceptual study of the three-way contrast in Polish. The acoustic results indicated that the center of gravity (COG) differences between the three fricatives was not particularly large: dental, 6445 Hz, retroflex, 5592 Hz; alveolopalatal, 5619 Hz. Different vocalic contexts resulted in the shifting of the spectral peak, rather than shifting the overall COG. They also found that the difference in vocalic transitions for F2 was significantly larger for the alveolopalatals than they were for both retroflex and dental segments. The F2 difference also persisted into the following vowel to a greater degree than the vowels following the other fricatives. Nowak [4] also performed a series of perceptual experiments to test the importance of the vocalic transitions and frication in the perception of Polish fricatives. Cross-splicing the vocalic transitions for alveolopalatals with retroflex frication resulted in a drop of accuracy below 80%. Furthermore, the accurate perception of alveolopalatals dropped significantly when the preceding vowel maintained vocalic transitions, but the following vowel had the transitions removed. Nowak [4] states that this implies the following vowel is more relevant for perceptual cues than the preceding one.

The purpose of the current study is two-fold: Lower Sorbian represents an understudied moribund language [5]; accurate accounts of the acoustic and articulatory characteristics of the Lower Sorbian inventory is sorely needed. Therefore, the first goal of this paper is to produce an accurate description of the three-way sibilant contrast in Lower Sorbian. The second goal of this paper is to identify the acoustic cues that listeners may use to distinguish the fricatives in a three-way contrast.

2. Methods

2.1. Participants

Two male (LS 1-2) and two female (LS 3-4) native speakers of Lower Sorbian participated in this study. The speakers Lower Sorbian speakers were also fluent in German through language contact. The speakers’ age ranged from 53-88 years of age, with a mean age of 73 years. The participants were recruited in the Sorbian speaking area around Cottbus, Germany. The participants had no self-reported speaking or hearing disorders.

2.2. Stimuli

The target phonemes, /s, z, ʂ, ʐ, ɕ, z/, were produced in word initial position using the real words: sarna ‘roe buck,’ zalza ‘gland,’ sapar ‘sheppard,’ żalość ‘pity,’ zasek ‘thin,’ and śatk ‘ten.’ Distracter tokens were also presented during elicitation. The participants repeated each word 3 times, for a total of 18
tokens from each participant and a total of 72 tokens (3 repetitions x 6 phonemes x 4 participants).

### 2.3. Instrumentation

Acoustic data for each participant was collected in a quiet room in each participant’s house using a Fostex FR-2 LE Field Recorder and a Lavalier AT831b microphone, recorded at 44,100 Hz and 32-bits.

### 2.4. Procedure and analysis

The Lavalier microphone was placed approximately 3 feet in front of the participant. Participants began by reading *The North Wind and the Sun* translated into Lower Sorbian two times, then they read the word list. The word list was comprised of the target words and distractor tokens. The tokens were all randomized and no experimental tokens were adjacent to each other. The participant read each token from the word list three times in a row and then proceed to the next word.

The fricatives were first filtered using a Hann pass filter with a minimum set at 1,000 Hz and a maximum set at 15,000 Hz. The purpose of the filter was to eliminate the energy from voicing so the voiced and voiceless pairs could be compared more reliably. Center of gravity (COG), skew, kurtosis and standard deviation (SD) measures were all taken with Praat [6] using a 30 ms window centered at the mid-point of the fricative. The spectral slopes for each fricative were calculated following the methods in [ygis, Pape, & Jesus [2]: Thomson multitaper power spectral density estimates [7] were created around the average peak frequency for all phonemes, F, which was 4350 Hz. The slopes M1 and M2 were then calculated with a best fit regression line using MATLAB [8]. F1 to F3 and F2-F1 [9] measures were also taken in Praat at the onset of each vowel using a 30 ms window and a 5,000 Hz ceiling. A repeated measures ANOVA was performed for each of the acoustic measures with the factor Fricative Identity (6 levels: s, z, ş, ʐ, ç, ʑ) in R [10] using the package ezANOVA [11]. The data was tested for normality using Mauchly’s Test for Sphericity and where the data was found not to be normal, the Greenhouse-Geisser corrected p-value and eigenvalue (GGe) are presented. Post-hoc pairwise t-tests were performed on the results for the kurtosis measure indicated no main effect of Fricative Identity after Greenhouse-Geisser corrections [F(5,15) = 7.14, GGe = 0.4029, p = 0.0136]. Post-hoc tests revealed that /s/ was higher than the retroflexes (/, ş, ʐ, /, p < 0.0001) and the alveolo-palatals (/ś, ʐ, p < 0.0001). /s, ʐ/ both had a higher skew than /z/ (p = 0.0026; p < 0.0001), /ś, /z/ had a higher skew than /ź/ (p < 0.0001), while only /ź/ had a lower skew than /ʐ/. See Table 1 and Figure 2 for a summary of the results.

### 3. Results

In this section, the acoustic measures for the fricatives are presented first (Section 3.1). This is followed by a presentation of the vowel transitions (Section 3.2).

#### 3.1. Acoustic measures of the fricatives

COG results indicated a main effect of Fricative Identity [F(5,15) = 53.89, p < 0.0001]. The post-hoc analysis indicated that /s/ had a higher frequency of friction than both of the retroflex segments (/ś, ʐ, p < 0.0001) and both of the alveolo-palatals (/ś, ʐ, p < 0.0001). However, /s/ was not found to be different from /ź/. The retroflex, /ś/, had a higher frequency than /ź/ (p = 0.0635) but a lower frequency than /ź/ (p < 0.0001). /ś/ had lower frequency than /ź/ and a higher frequency than /ź/. However, it was not found to be significantly different from /ś, ʐ/. See Table 1 and Figure 1 for a summary of the COG results.
The results for the analysis of the slope, M1, revealed a significant effect of Fricative Identity \([F(5,15) = 11.41, p = 0.0001]\). The post-hoc comparisons indicated that /\(s/\) had a lower slope than /\(z/\) (\(p < 0.0001\)), both the alveolopalatals (/\(z/\), \(p < 0.0002\); /\(z/\), \(p < 0.0001\)), and /\(z/\) (\(p = 0.0079\)), but not /\(z/\) (\(p = 1.0000\)). /\(z/\) had a lower slope than both /\(s/\) (\(p < 0.0001\)) and /\(z/\) (\(p < 0.0001\)), but a higher slope than /\(z/\) (\(p < 0.0001\)). /\(z/\) had a lower slope than /\(z/\) (\(p < 0.0001\)) and /\(z/\) (\(p < 0.0001\)), but a higher slope than /\(z/\) (\(p = 1.0000\)) or /\(z/\) (\(p = 1.0000\)), but it was found to be higher than /\(z/\) (\(p = 0.00050\)). See Table 2 for a summary of the results.

The results for the analysis of the slope M2 revealed a main effect of Fricative Identity \([F(5,15) = 4.66, p = 0.0090]\). The post-hoc analysis revealed that the slope, M1, was lower for /\(s, z/\) than /\(ʂ/\) (\(p = 0.0022\), \(p = 0.0027\)), but not /\(ʐ/\) (\(p = 0.0819\); \(p = 0.0967\)). /\(s, z/\) were also found to have a lower slope than /\(ɕ/\) (\(p = 0.0464\); \(p = 0.0552\)), but not /\(ʑ/\) (\(p = 0.0797\), \(p = 0.0941\)). /\(s/\) and /\(z/\) did not differ from each other (\(p = 1.0000\)). The slope, M2, did not reliably distinguish the retroflexes from the alveolopalatals for any comparison (\(p = 1.0000\) in all cases. See Table 2 for a summary of the results.

### Table 2. Summary of the mean slopes M1 and M2.

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>/s/</td>
<td>-0.0017</td>
<td>-0.0007</td>
</tr>
<tr>
<td>/z/</td>
<td>-0.0045</td>
<td>-0.0007</td>
</tr>
<tr>
<td>/ʂ/</td>
<td>0.0032</td>
<td>-0.0033</td>
</tr>
<tr>
<td>/ʐ/</td>
<td>-0.0015</td>
<td>-0.0031</td>
</tr>
<tr>
<td>/ɕ/</td>
<td>0.0031</td>
<td>-0.0041</td>
</tr>
<tr>
<td>/ʑ/</td>
<td>-0.0045</td>
<td>-0.0031</td>
</tr>
</tbody>
</table>

### 3.2. Vowel transitions

The results for the analysis of F1 revealed a main effect of Fricative Identity \([F(5,15) = 17.84, p < 0.0001]\). The post-hoc analysis indicated that /\(s/\) had a higher F1 than /\(z/\) (\(p = 0.0001\), /\(z/\) (\(p < 0.0001\), /\(z/\) (\(p = 0.0437\), and /\(z/\) (\(p = 0.0001\)), but no significant difference from /\(z/\) (\(p = 1.0000\)). /\(z/\) had a lower F1 than /\(z/\) (\(p < 0.0001\), but no significant difference from /\(z/\) or the alveolopalatals (\(p = 1.0000\) in all cases). /\(z/\) had a higher F1 than all of the voiced fricatives /\(z, z, z/\) (\(p < 0.0001\) in all cases). /\(z/\) had a lower F1 than /\(z/\) (\(p = 0.0093\)), but a higher F1 than /\(z/\) (\(p = 0.0339\)). There was no significant difference between /\(z/\) and /\(z, z, z/\) (\(p = 1.0000\)). See Table 3 and Figure 4 for a summary of the F1 results.

The F2 results revealed a main effect of Fricative Identity \([F(5,15) = 5.04, p = 0.0066]\). Post-hoc tests revealed that /\(s/\) had a lower F2 than /\(z/\) (\(p = 0.0475\) and marginally lower F2 than /\(z/\) (\(p = 0.0752\)). However, there was no significant difference in the F2 transitions for /\(s/\) and /\(ʂ, z, z/\) had a lower F2 than the retroflexes (\(p < 0.0001\) and /\(z/\) (\(p = 0.0207\) but not /\(z/\) (\(p = 0.6891\)). /\(z/\) had a higher F2 than /\(z/\) (\(p = 0.0047\), but not /\(z/\) (\(p = 0.2246\) or /\(z/\) (\(p = 1.0000\). /\(z/\) was also found to have a higher F2 than /\(z/\) (\(p = 0.0079\), but not /\(z/\) (\(p = 0.3359\)). Finally, /\(z/\) showed no significant difference from /\(z/\) (\(p = 1.0000\)). Table 3 and Figure 5 summarizes the F2 results. The F3 results indicated no main effect of Fricative Identity \([F(5,15) = 1.47, p = 0.2556]\). See Table 3 for a summary of the F3 results.

The F2-F1 results indicated a main effect of fricative identity \([F(5,15) = 11.44, p < 0.0001]\). Post-hoc comparisons revealed the alveolars, /\(s, z/\) had a lower F2-F1 than both alveolopalatals segments (\(p < 0.0001\), but not the retroflexes (\(z/\) (\(p = 1.0000\); /\(z/\) (\(p = 0.1670\) or /\(z/\) (\(p = 1.0000\)). The alveolopalatals, /\(s, z/\), had a higher F2-F1 than both the alveolars (\(p < 0.0001\) and both of the retroflexes (\(z/\) (\(p < 0.0001\); /\(z/\) (\(p = 0.0310\)). See Table 3 and Figure 6 for a summary of the results.

### Table 3. Summary of mean slopes of F1-F3.

<table>
<thead>
<tr>
<th>M1</th>
<th>M2</th>
</tr>
</thead>
<tbody>
<tr>
<td>/s/</td>
<td>-0.0017</td>
</tr>
<tr>
<td>/z/</td>
<td>-0.0045</td>
</tr>
<tr>
<td>/ʂ/</td>
<td>0.0032</td>
</tr>
<tr>
<td>/ʐ/</td>
<td>-0.0015</td>
</tr>
<tr>
<td>/ɕ/</td>
<td>0.0031</td>
</tr>
<tr>
<td>/ʑ/</td>
<td>-0.0045</td>
</tr>
</tbody>
</table>

Figure 3: COG, Skew, and SD for each of the fricatives.

Figure 4: Boxplot of the F1 results.

Figure 5: Boxplot of the F2 results.

Figure 6: Boxplot of the F2-F1 results.
Table 3. Summary of the mean F1 and F2 transitions (Hz).

| s  | 746 | 1438 | 2478 | 692 |
| z  | 528 | 1234 | 2296 | 706 |
| s̝ | 768 | 1480 | 2480 | 1056 |
| z̝ | 528 | 1384 | 2213 | 1186 |
| ṣ | 608 | 1664 | 2521 | 711 |
| s̢ | 466 | 1652 | 2536 | 855 |

4. Discussion

An acoustic analysis of the three-way contrast in Lower Sorbian has been presented. Acoustic measures of the fricative spectra and the vocalic transitions have been presented.

The overall results suggest that there is a three-way contrast present in Lower Sorbian, which is best described by a series of alveolars, retroflexes and alveolopalatals. The clearest indication of this is through the examination of the COG values in conjunction with the F1 and F2 values. The COG for the alveolars clearly distinguishes it from the other two fricative series. The F2 measures for the alveolars and the retroflex segments indicate the place of articulation is in a similar place. However, the COGs show a drastic difference in the frication. This is likely a result of the overall tongue shape: the alveolars have a deep groove and a forward place of articulation; the retroflex segments have a slightly drawn up tongue tip, with a flatter cross-sectional contour. However, the drawn up tongue tip for the retroflexes creates a sub-apical cavity and likely impedes airflow more than the mid-line groove for the alveolars. Ultimately, this suggests that overall tongue shape plays a large role in the COG of a fricative segment, not just the length of the front cavity.

The COG values for the retroflex and alveolopalatals did not differ significantly from each other. However, there was a clear distinction in both F1 and F2 for the alveolopalatals, suggesting they are articulated both higher and further forward in the mouth. In fact, the F2 results suggest that the tongue-palate contact extends as far as the dental place of articulation. However, the low COG suggest the constriction extends much further back in the mouth. These findings also suggest that the overall length of the constriction greatly affects the frequency for the COG. Even though the front cavity for the alveolopalatals is relatively short - possibly the shortest of all three sibilant series - it has a relatively low COG.

Another interesting discovery in this paper is that the three-way contrast is not realized in the same way as in Polish. Nowak [4] reports a much higher COG for the retroflex (5592 Hz) and alveolopalatal (5619 Hz) segments in Polish. However, the mean COG of the retroflex segments in Lower Sorbian was 3777 Hz for /ʂ/ and 2472 Hz for /ʐ/. The COG for the alveolopalatal segments was not found to distinguish them from each other. This suggests that COG, in conjunction with vowel transitions, plays a major role in distinguishing these two segments. This is in contrast to the findings in Nowak [4], which indicated that either vowel transitions on the following vowel or the fricative spectra alone was sufficient for distinguishing the sibilants in Polish. Ultimately, the data presented here could suggest that the three-way contrast is not composed of the same segments. It is clear that the alveolopalatals and retroflex segments have a significantly lower COG (approximately 2,000 Hz). The F1 and F2 measures presented suggest that the second segment is a retroflex. This follows from Keating's [13] assertion that the first two formants for retroflex fricatives are typically found in approximately the 1,600-2,400 Hz range. The formant transitions also suggest an alveolopalatal for the third segment due to a low F1 and higher F2 (in the 600-1600 Hz range). It remains unclear what the tongue configuration difference is between the Lower Sorbian and Polish sibilants, but an articulatory study is required to validate these claims.

5. Acknowledgments

I would like to thank Dr. Kaufirst, the Sorbian Institute in Chöšebuz, and Ingolf Kschenka. This work was funded by the Germany/Europe Fund granted by the University of Toronto.

6. References

