GENDER-SPECIFIC DIFFERENCES IN SIBILANT CONTRAST REALIZATIONS IN ENGLISH AND GERMAN

Melanie Weirich & Adrian P. Simpson
Institute of German Linguistics, Friedrich Schiller University Jena, Germany
melanie.weirich@uni-jena.de adrian.simpson@uni-jena.de

ABSTRACT

This study examines possible differences between the acoustic realization of the intersibilant contrast /s/~/ʃ/ in German and American English. A range of acoustic parameters (COG, standard deviation, skewness, kurtosis and Discrete Cosine Transformation coefficients) are calculated to characterize the spectra of the two sibilants. Significant differences are found between the male and female intersibilant contrast, indicating that females produce a stronger acoustic contrast between /s/ and /ʃ/ in both languages. While in the German data set a tendency for gender-specific differences in accent realizations was found, the effect did not reach significance.

Keywords: gender-specific differences, sibilants, DCT, spectral moments

1. INTRODUCTION

Besides lexical information, phonetic detail also transmits indexical information about the speaker [5]. Such idiosyncratic features might have a biological origin or be the product of learned speech behaviour [14]. Sibilants, and in particular /s/, have been claimed to potentially index both sex (i.e. physical inevitabilities such as anatomical differences of vocal tract size) and gender (i.e. performed, constructed, learned gender identities) [6, 18, 20]. Females have been found to show /s/ productions with higher spectral energy peaks than males [3, 13]. One possible line of explanation is based on potential anatomical differences in the speech apparatus between speakers/sexes [6, 18, 19]. Weirich & Fuchs [23] found a relationship between the alveolo-palatal steepness of a speaker and his/her articulatory realization of the sibilant contrast /s/-/ʃ/ in German. However, due to the relatively small subject group examined in that study, no comparison between males and females could be made. Besides organic factors, learned behavioural sources have also been identified, such as socio-cultural background or the conscious choice of an indexical feature to display a certain social identity or group membership [20]. We shall use the term *gender* throughout this paper, fully aware that both anatomical inevitabilities, as well as learned behavioural aspects are involved in the patterns we describe.

Higher spectral energy in /s/, typically found in female speech, often goes hand in hand with a larger acoustic distance to /ʃ/. Romeo et al. [17] found not only a larger acoustic distance between /s/ and /ʃ/ in females than in males, but also a larger contrast between /p/ and /b/ in females. This is in line with the larger acoustic vowel space that has been repeatedly reported for female speakers [1, 10, 24]. All of these findings seem to indicate that female speakers are realizing stronger phonological contrasts than males, suggesting that they are speaking more clearly [9], even though the reasons why female speakers should be speaking more clearly are not so apparent.

In this study, we investigate the acoustic correlates of /s/ and /ʃ/ in two languages: German and American English. We will focus in particular on potential gender-specific differences in the realization of the phoneme contrast, at the same time examining any cross-linguistic similarities. An acoustic parameterization of the sibilant spectra is carried out using the four spectral moments (COG, SD, kurtosis and skewness) and additionally Discrete Cosine Transformation (DCT) [22].

2. METHOD

2.1. Subjects and speech material

The study comprises two sets of data. The American English data set consists of a subset of the University of Wisconsin X-Ray Microbeam Speech Production Database [26] and comprises 34 speakers (20 f, 14 m). The sibilants /s/ and /ʃ/ were analysed in a list containing several aCa-sequences, among them /asa/ and /aʃa/. For each speaker one repetition of each sibilant was used.

The German data set comprises 11 speakers (6f, 5m). Acoustic recordings were made at Potsdam University as part of a larger study examining gender-related articulatory and acoustic differences. Contained within a larger corpus of stimuli designed, among other things, to elicit different prosodic structures were the sentences *Die blaue Tasse/Tasche ist im Schrank* (‘The blue cup/bag is in the cupboard’), also containing the target sibilants /s/ and /ʃ/ in intervocalic position. In addition to this control condition, the sentences were recorded in two different accent conditions by using a question-answer paradigm with accents either on the target word *Tasse/Tasche* (accented condition) or the preceding adjective *blaue* (unaccented condition).

2.2. Acoustic analysis

For the acoustic characterization of the sibilants several parameters were calculated that have been found to play a role in the acoustics of sibilants (e.g. [4, 11, 13, 15]). The first four spectral moments [4] were calculated in PRAAT: 1) centroid/*Center of Gravity* (COG), 2) *variance* or *standard deviation*, 3) *skewness*, 4) *kurtosis*. All measurements were made around the acoustic midpoint of the sibilant with a window length of 0.025s and a cutoff frequency of 500 Hz to control for potentially voiced segments.
Second, we used *Discrete Cosine Transformation* (DCT), a method proposed by Watson & Harrington [22], to parameterize the shape of the spectra and, in particular, to quantify the acoustic contrast between sibilants. This method decomposes the signal into a set of half-cycle cosine waves and the resulting amplitudes of these cosine waves are the DCT coefficients (corresponding to the cepstral coefficients of a spectrum). Three DCT coefficients were used for the analysis. DCT1 is proportional to the linear slope of the spectrum, DCT2 corresponds to its curvature and DCT3 describes the amplitude of the higher frequencies. Guzik & Harrington [7] showed that the DCTs provide a very effective separation between the four fricative types in Polish, and Jannedy et al. [12] found DCTs to be a reliable parameter to differentiate the very similar acoustic spectra of /ç/ and /ʃ/ in Berlin German. DCT transformation was applied after the spectra had been converted to the Bark scale following [8].

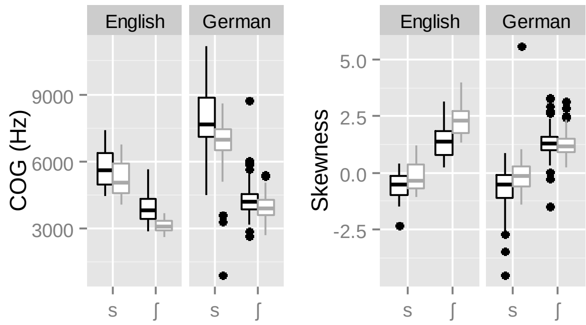
3. RESULTS

All statistical analyses were carried out using the R environment. Linear mixed models were run with the lme4 package and likelihood ratio tests were used to find the model with the best fit to the data and determine significant main effects and/or interactions.

3.1 Acoustics of sibilants: spectral moments

For the English data we found a significant main effect of sibilant class and gender for COG, kurtosis and skewness. For COG, higher values were found for /s/ than for /ʃ/ (Estimate: 1934Hz, Standard error 121.1 Hz, $p < .001$) and for females than for males (Estimate: 706.2 Hz, Standard error 199.8 Hz, $p < .01$, see Figure 1). For skewness, /ʃ/ showed higher values than /s/ (Estimate: 2.3, SE: 0.14, $p < .01$) and males showed higher values than females (Estimate: 0.72, SE: 0.19, p < .05). For kurtosis, higher values were found for /ʃ/ than for /s/ (Estimate: 7.6, SE: 1.2, p < .01) and higher values were found for males than for females (Estimate: 2.9, SE: 1.3, $p < .05$). For SD we found a significant interaction of gender*sibilant class indicating that females have a significantly higher value (for about 500Hz, $p < .01$) than males only for /s/.

Figure 1: COG and skewness values plotted as a function of gender (f: black, m: grey), sibilant and language

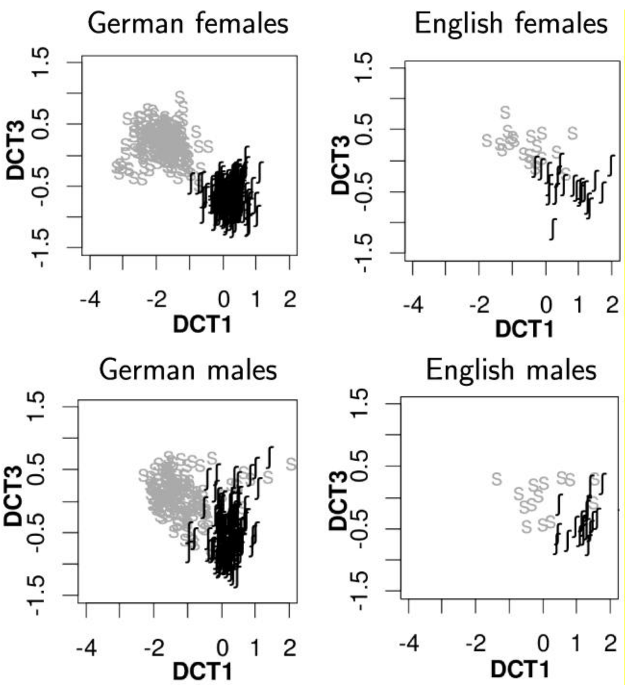


For the German data we found a significant effect of gender in terms of an interaction with the sibilant class for COG and skewness (see Figure 1). In both cases, men and women differed significantly only for /s/ but not for /ʃ/: females exhibit higher COG values than males (for about 1079 Hz, Standard error = 264.43 Hz, $p < .05$) and more negative skewness (for about 0.53, standard error = 0.18, $p < .05$). As Figure 1 shows, the difference between the genders in /s/ strengthens the contrast between the sibilants for females. No interaction of gender and accent was found.
To summarize, in both languages significant effects of speaker gender were found. However, they were found more often for English than for German, possibly reflecting the differences in sample sizes. Also, if an interaction with sibilant class was found, it indicated a higher sensitivity for /s/ showing gender differences.

3.2 Acoustic realization of the sibilant contrast: DCTs

As mentioned above the estimated DCT coefficients were used to describe and quantify further the acoustic distance between the sibilants. Figure 2 shows the third DCT-coefficient (distribution of energy in high frequencies) plotted as a function of the first coefficient (linear slope) for /s/ and /ʃ/, calculated from Bark-scaled spectra. The plots are separated by gender (above: females, below: males) and language (left: German, right: English). It appears that – as expected – females spaced their sibilants further apart acoustically than males. This seems to be based particularly on differences in the /s/ productions.

Figure 2: DCT3 as a function of DCT1 separated by language, gender and sibilant



For the German data and DCT1, 2 and 3 a significant interaction of sibilant class and gender was found: While females and males did not differ in DCT1 for /ʃ/, females have significantly higher DCT1 values than males for /s/ (Estimate: 0.55, SE: 0.12, $p < .05$). For DCT2, only females showed a significant difference between the sibilants (Estimate: 0.21, SE: 0.04, $p < .05$.). For DCT3 the significant interaction indicated that while both genders differ between the sibilants in DCT3, the difference is significantly larger in females than in males (Estimate: 0.24, SE: 0.04, $p < .05$).

For the English data we found significant main effects for gender and sibilant class for DCT1, with higher values for males than females (Estimate: 0.55, SE: 0.17, $p < .05$), and /ʃ/ than /s/ (Estimate: 1.22, SE: 0.12, p < .01). For DCT3, the same interaction between gender and sibilant class as for the German data was found: the difference between the sounds was significantly larger for females than for males (Estimate: 0.31, SE: 0.09, $p < .05$).

For a better quantification of the acoustic contrast between the sibilants, Euclidean Distances (EDs) in the DCT1xDCT2xDCT3 space were calculated. For both languages we found a comparable main effect of gender, with females having a larger acoustic contrast than males (German = Estimate: 0.64, SE: 0.25, $p < .05$, English = Estimate: 0.65, SE: 0.30, $p < .05$). While the strength of the gender-specific difference is comparable between the languages, the contrast is generally larger (for both genders) in the German data than in the English data (Estimate: 0.70, SE: 0.19, $p < .05$)

Figure 3: EDs in DCT1xDCT2xDCT3 space separated by gender and language

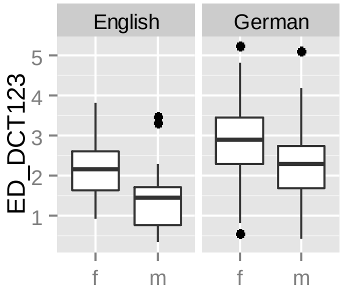
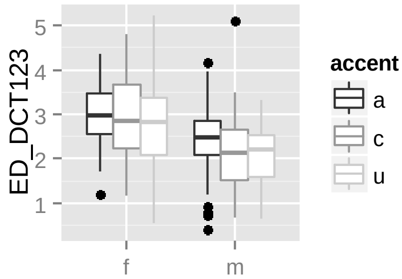


Figure 4 shows the acoustic contrast of the German data separated by the additional factor accent condition. Even though the figure points to a stronger effect of accent condition on males than on females, no significant effect of accent or an interaction of gender*accent was found.

Figure 4: EDs in DCT1xDCT2xDCT3 space separated by gender and accent condition



4. DISCUSSION

This study examined possible gender-related differences in the acoustic realization of the sibilant contrast in German and American English. Despite differences in sample size (34 vs. 11 speakers) and elicitation (meaningless aCa sequence vs. real word), male and female speakers exhibit similar differences across the two languages, with female speakers, as expected from the findings of earlier studies, e.g. [17], realizing a stronger acoustic contrast than the male speakers. In the German data set we were also able to analyse possible differences relating to different accent conditions. In line with studies that found females producing larger durational differences between vowel categories in accented conditions [21, 2], we might have expected females to enhance the intersibilant contrast under accent even further. This does not seem to be the case. While the EDs plotted in figure 4 indicate that the female intersibilant contrast is stronger than the male, there is no difference between the different accent conditions. Indeed, if a difference is present between the different accent conditions, it would seem to be being made by the male speakers, although this does not reach significance. A similar tendency was found for the realization of the vowel length contrast in German [25]: here, an interaction of gender and accent condition was found with males showing a larger difference between accented and unaccented condition than females. Reasons for a lack of difference are unclear, but it is possible that the articulatory precision required for sibilants, in turn reduces the space for acoustic flexibility [16].

We are currently examining the articulatory differences giving rise to the acoustic differences we have described here. The results of the present study indicate that differences in the acoustics of intersibilant contrast are part of the learned indexicalization of gender. The next step will be to investigate the articulatory origins of the observed acoustic differences. This may cast more light on whether the acoustic differences are indeed produced by significant articulatory differences. Alternatively, as has also been found for vowels, similar articulatory differences between sibilants might conceivably be giving rise to significantly different acoustic products.

5. REFERENCES

[1] Diehl, R. L., Lindblom, B., Hoemeke, K. A. & Fahey, R. P. 1996. On explaining certain male–female differences in the phonetic realization of vowel categories, *J. Phonetics* 24, 187–208.

[2] Ericsdotter, C. & Ericsson, A. M. 2001. Gender differences in vowel duration in read Swedish: preliminary results. *Proc. Fonetik* Lund, 34–37.

[3] Flipsen, P.; Shriberg, L.; Weismer, G.; Karlsson, H. & McSweeney, J. 1999. Acoustic characteristics of /s/ in adolescents. *J. Speech Lang. Hear. R.* 423, 663-677.

[4] Forrest, K.; Weismer, G.; Milenkovic, P. & Dougall, R. N. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *J. Acoust. Soc. Am.* 841, 115-123.

[5] Foulkes, P. & Docherty, G. 2006. The social life of phonetics and phonology. *J Phonetics* 34, 409–438.

[6] Fuchs, S. & Toda, M. 2010. Do differences in male versus female /s/ reflect biological or sociophonetic factors? In S. Fuchs, M. Toda,& M. Zygis Eds., *Turbulent sounds: An interdisciplinary guide,* Mouton de Gruyter, Berlin, Germany, pp. 281–302.

[7] Guzik, K. & Harrington, J. 2007. The quantification of place of articulation assimilation in electropalatographic data using the similarity index SI. *Advances in Speech Language Pathology* 91, 109-119.

[8] Harrington, J. 2010. *Phonetic analysis of speech corpora*. Wiley-Blackwell.

[9]Henton, C. G. 1995. Cross-language variation in the vowels of female and male speakers, *Proc. XIIIth ICPhS* vol. 4, Stockholm, 420–423.

[10] Hillenbrand, J., L. A. Getty, M. J. Clark, and K. Wheeler 1995. Acoustic characteristics of American English vowels, *J. Acoust. Soc. Am.* 97, 3099–3111.

[11] Hughes, G. & Halle, M. 1956 Spectral properties of fricative consonants. *J. Acoust. Soc. Am.* 28, 303-310.

[12] Jannedy, S.; Weirich, M.; Brunner, J. & Mertins, M. 2010 Perceptual evidence for allophonic variation of the palatal fricative /ç/ in spontaneous Berlin German. *J. Acoust. Soc. Am.* 128, 2458.

[13] Jongman, A.; Wayland, R. & Wong, S. 2000. Acoustic characteristics of English fricatives. *J. Acoust. Soc. Am.* 108, 1252-1263.

[14] Ladefoged, P. & Broadbent, D. 1957. Information conveyed by vowels. *J. Acoust. Soc. Am.* 29, 98–104.

[15] Newman, R. S. 2003. Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report. *J. Acoust. Soc. Am.* 1135, 2850-2860.

[16] Nicolaidis, K. 2001. An electropalatographic study of Greek spontaneous speech. *Journal of the International Phonetic Association* 31, 67–85.

[17] Romeo, R., Hazan, V., Pettinato, M. 2013. Developmental and gender-related trends of intra-talker variability in consonant production, *J. Acoust. Soc. Am.* 134 5, 3781-3792.

[18] Rudy, K., & Yunusova, Y. 2013. The effect of anatomic factors on tongue position variability during consonants. *J. Speech Lang. Hear. R.*, 56, 137–149.

[19] Stevens, K. N. 1998. *Acoustic Phonetics*. MIT Press.

[20] Stuart-Smith, J. 2007. Empirical evidence for gendered speech production: /s/ in Glaswegian. In: Cole, J. and Hualde, J.I. eds. *Laboratory Phonology* 9. Series: Phonology and phonetics. Mouton de Gruyter, New York, USA, pp. 65-86.

[21] Wassink, A. B. 1999. *A sociophonetic analysis of Jamaican vowels.* Ph.D. thesis, University of Michigan, Michigan.

[22] Watson, C. I. & Harrington, J. 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *J. Acoust. Soc. Am.* 106, 458-468.

[23] Weirich, M. & Fuchs, S. 2013. Palatal morphology can influence speaker-specific realizations of phonemic contrasts. *J. Speech Lang. Hear. R.* 56, S1894-S1908.

[24] Weirich, M. & Simpson, A. 2014. Differences in acoustic vowel space and the perception of speech tempo, *J. Phonetics* 43. 1-10.

[25] Weirich, M. & Simpson, A.P. In press. Impact and interaction of accent realization and speaker's sex on vowel length in German. In: A. Leemann, M.-J. Kolly, S. Schmid & V. Dellwo eds., *Trends in Phonetics and Phonology in German speaking Europe.* Frankfurt/M.: Peter Lang.

[26] Westbury, J. R. 1994. *X-ray Microbeam Speech Production Database User's Handbook Version 1.0.* Waisman Center on Mental Retardation and Human Development, Madison, WI.