Do differences in male versus female /s/ reflect biological or sociophonetic factors?

Susanne Fuchs and Martine Toda

1. Theoretical background

For many decades researchers have been preoccupied with the search for invariance at the acoustic, articulatory and motor control levels which could potentially correspond to phonological entities. This undertaking has more recently been redirected towards the discussion of the extent of variability and towards the question of what causes variability (Perkell et al. 2004, Newman, Clouse and Burnham 2001). Explanations in the literature are manifold: variability can be a consequence of (a) phonemic inventory; (b) phonemic context; (c) stress, speech rate or loudness; (d) can be driven by perceptual requirements; (e) caused by the position of the segment in the syllable, word and other constituents; (f) can be a result of the dialectal origin of the speaker, the speaker’s gender, the speaker’s vocal tract, age or mood, among many other possibilities.

For the purposes of the current study we will differentiate between biological and non-biological explanations. In our view, biological factors are those that cause variability due to the bio-physical properties of the speech production and perception systems and changes in these properties during ontogenesis (Fuchs, Pomino-Marshall and Perrier 2007). For example, differences in fundamental frequency between young children and male adults should to a great extent be attributed to differences in the size of the vocal folds and their respective mass and tension, i.e. properties that are anatomically given.

In contrast, non-biological factors arise from speech and language as a collective behaviour which is learnt during speech acquisition and is both adaptable and robust during ontogenesis (Pierrehumbert 2006). So far empirical evidence from socially conditioned variability comes mainly from the areas of sociolinguistics, sociology, and psychology. However, as phonetics has become more and more interdisciplinary, researchers have also devoted increasing attention to this topic, for example in a special issue of the Journal of Phonetics (edited by Jannedy and Hay 2006) on socially
conditioned variability in speech production and perception. Pierrehumbert and Clopper (forthcoming) even call sociophonetic issues the next challenge for Laboratory Phonology.

Our present work focuses on the potential biological versus sociophonetic origin of phonetic differences in male versus female /s/ production, an issue which has previously been raised on the basis of acoustic and perceptual data. We adapt the terminology used by Stuart-Smith (2007) and many others and call male-female differences sex differences when they are based on biological factors. The term gender will be used instead whenever differences are grounded in sociophonetic issues. In other words:

“Gender is described as what we ‘do’, or what we ‘perform’: gender doesn’t just exist, but is continually produced, reproduced, and indeed changed though people’s performance of gendered acts, as they project their own claimed gender identities, ratify or challenge others’ identities, and in various ways support or challenge systems of gender relations and privilege “(Stuart-Smith 2007, p. 66 citing Eckert and McConnell-Ginet).

1.1. Acoustic and perceptual evidence for differences in male and female /s/ production

At this point we would like to provide some examples from the literature of differences in male versus female /s/ realization. The earliest work we are aware of comes from the perceptual domain. Schwartz (1968) studied listeners’ ability to identify speakers’ sex from a variety of isolated voiceless fricatives. Results of his study showed that listeners were able to identify speakers’ sex for the sibilants, but not for /f, θ/. This was attributed to the higher frequencies in females’ realizations in comparison to males’ realizations of the sibilants. Schwartz supposed that differences are based on vocal tract differences. Similarly, Johnson (1991) discussed his perceptual findings on speaker and vowel variability in a /s/ to /ʃ/ continuum with respect to biological factors, i.e., to anatomical differences between males and females.

A detailed literature review of studies of /s/ is provided in Flipsen et al. (1999). In their own acoustic investigation with 26 American English speakers between 9 and 15 years of age they found significant sex-related spectral differences, in particular in the frequency mean and skewness parameter at the fricative midpoint.
Gordon, Barthmaier and Sands (2002) investigated fricative inventories in seven genetically unrelated languages. They used the acoustic centre of gravity parameter and found male-female differences in Chikasaw. The authors reported no differences in the other languages, but this may reflect small sample sizes and that the authors grouped several fricatives together.

Additional evidence comes from an acoustic study by Heffernan (2004) which included 12 Canadian English speakers and 10 Japanese speakers. He found significant male-female differences for the acoustic centre of gravity parameter. The male-female distinction was generally more pronounced in the Canadian English speakers than in the Japanese speakers. He interpreted these results with respect to a social component to explain male-female differences, although a potential mixture with physiological differences was discussed too.

1.2. How can male-female differences in /s/ be explained?

The question which motivated our own work is: Are the acoustic differences in males’ and females’ realization of /s/ a consequence of differences in vocal tract size, or are they a consequence of a sociophonetic process in which females actively produce /s/ with different articulatory strategies than males as an index of their gender? The former explanation relates to the biological nature of speech underlying interspeaker variation whereas the latter is grounded in the social function of speech. A combination of the two explanations may be possible as well.

Support for the biological explanation includes the fact that acoustic male-female differences are found in various genetically unrelated languages (e.g., English, Japanese, and Chickasaw) and they always go in the same direction, with females realizing /s/ with higher frequencies than males. Stevens (1998, p. 398) also notes that the length of the cavity downstream of the oral constriction is somewhat smaller for females than for males. Since small differences in the size of the front cavity, especially in its length, are crucial for the spectral properties of /s/, sex differences can be expected. Another argument in favor of the biological explanation would be that acoustic differences are often found in sibilants where vocal tract differences could matter, but not in labiodentals, where vocal tract differences do not matter.

Support for the gender explanation comes from sociophonetically grounded work (Stuart-Smith 2007, Strand 1999, Heffernan 2004, Munson
et al. 2006). One of the main arguments of these studies is that morphological differences between males and females are larger in the back part of the vocal tract, i.e. in the pharynx (see Fitch and Giedd 1999 for a very comprehensive study), than at the place where alveolar fricatives are realized. Another piece of evidence comes from experiments including males and females of different ages and social classes. Stuart-Smith (2007), for instance, reports for Glaswegian English that although there may be vocal tract differences between younger (13–14 years) and older (40–60 years) males, the measured acoustic parameters were relatively homogeneous. By contrast, the young working class girls behaved similarly to all males, but differed from all other females (older working class and middle class women, middle class girls).

Another recent study from Munson et al. (2006) supports the importance of sociophonetic factors in /s/ production. They investigated the acoustic and perceptual bases of women’s and men’s sexual orientation by means of homosexuals’, bisexuals’ and heterosexuals’ speech. Their acoustic findings provide evidence that homosexual and heterosexual men’s speech differs significantly in the spectral skewness parameter of /s/. (Results for lesbian and heterosexual females, however, did not reveal such patterns.) In a second experiment Munson et al. were able to show that listeners’ judgements of perceived sexual orientation were related to the acoustic findings.

Stuart-Smith’s and Munson et al.’s findings suggest a primary role of gender identity in the realization of /s/ and refute a primary role of biological causes. None of the studies we are aware of obtained direct morphological or articulatory data from their subjects.

The current study is a first attempt to close this gap by means of electropalatographic (EPG) data simultaneously recorded with acoustics. Additionally, we gathered morphological data for the size of the speaker’s palate from the electrodes located in the EPG palate. The main goal of this study is to determine whether male versus female /s/ realization is best accounted for by biological or by sociophonetic factors. In the next section we will outline how this question was addressed.

2. Methodology

The corpus we used was originally constructed to study the influence of the palate shape on token-to-token variability (Brunner, Fuchs and Perrier
Here we will only consider word-medial /s/ in the target words /sasa/ for the English speakers and /zasa/ for the Germans. Target words were embedded in the carrier phrase ‘Say ___ please.’ for the English subjects and ‘Habe ___ gesagt.’ (Have .... said.) for the German subjects. The target sibilants occurred word medially in an ambisyllabic post-stressed position. All sentences were presented in a randomized order and repeated thirty times (with some exceptions cf. table 1).

Table 1. Participants of the study with their language and gender abbreviation (E = English, G = German, M = male, F = female) and number of repetitions (N).

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Language (regional origin)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM1</td>
<td>English (Australian English)</td>
<td>30</td>
</tr>
<tr>
<td>EM2</td>
<td>English (American English)</td>
<td>30</td>
</tr>
<tr>
<td>EM3</td>
<td>English (Scottish English)</td>
<td>30</td>
</tr>
<tr>
<td>EM4</td>
<td>English (RP)</td>
<td>29</td>
</tr>
<tr>
<td>EM5</td>
<td>English (RP)</td>
<td>28</td>
</tr>
<tr>
<td>EM6</td>
<td>English (Scottish English)</td>
<td>30</td>
</tr>
<tr>
<td>EF1</td>
<td>English (RP)</td>
<td>30</td>
</tr>
<tr>
<td>EF2</td>
<td>English (Scottish English)</td>
<td>29</td>
</tr>
<tr>
<td>EF3</td>
<td>English (RP)</td>
<td>30</td>
</tr>
<tr>
<td>EF4</td>
<td>English (RP)</td>
<td>24</td>
</tr>
<tr>
<td>EF5</td>
<td>English (RP)</td>
<td>29</td>
</tr>
<tr>
<td>EF6</td>
<td>English (RP)</td>
<td>30</td>
</tr>
<tr>
<td>GM1</td>
<td>German (Bavarian)</td>
<td>30</td>
</tr>
<tr>
<td>GM2</td>
<td>German (Alemannic)</td>
<td>30</td>
</tr>
<tr>
<td>GM3</td>
<td>German (Northern German)</td>
<td>30</td>
</tr>
<tr>
<td>GM4</td>
<td>German (Northern German)</td>
<td>30</td>
</tr>
<tr>
<td>GM5</td>
<td>German (Bavarian)</td>
<td>30</td>
</tr>
<tr>
<td>GM6</td>
<td>German (Northern German)</td>
<td>25</td>
</tr>
<tr>
<td>GF1</td>
<td>German (Bavarian)</td>
<td>23</td>
</tr>
<tr>
<td>GF2</td>
<td>German (Northern German)</td>
<td>30</td>
</tr>
<tr>
<td>GF3</td>
<td>German (Northern German)</td>
<td>31</td>
</tr>
<tr>
<td>GF4</td>
<td>German (Northern German)</td>
<td>30</td>
</tr>
<tr>
<td>GF5</td>
<td>German (Saxonian)</td>
<td>30</td>
</tr>
<tr>
<td>GF6</td>
<td>German (Northern German)</td>
<td>31</td>
</tr>
<tr>
<td>Σ 24</td>
<td>English: 12 speakers</td>
<td>699</td>
</tr>
<tr>
<td></td>
<td>German: 12 speakers</td>
<td></td>
</tr>
</tbody>
</table>

Recordings of the German-speaking subjects took place in a professional soundproofed room at Zentrum für Allgemeine Sprachwissenschaft (ZAS)
in Berlin using the EPG 3 Reading system, and the English-speaking subjects were recorded at Queen Margaret University College (QMUC) Edinburgh using the Win EPG system; all recordings, however, involved the same kind of electrode arrangement on the artificial palate of the EPG 3 system. All EPG data have a sampling frequency of 100 Hz. Acoustic data were recorded with a sampling frequency of 44 kHz for the English speaking subjects.

For the German speaking subjects (except GM2, GM4, GF3, and GF6) and one English speaker (EM4) data were recorded with 48 kHz, but they were further downsampled to 16 kHz since the corpus was originally not recorded to study sibilants. The acoustic data with the original sampling rate are no longer available. Altogether 24 speakers were recorded (4 more speakers than the original corpus), 6 male and 6 female native speakers each of German and English. Since our experimental set-up restricted us to those subjects who had an EPG palate available, our sample is not equally well balanced with respect to age (ranging from the early 20s to mid 50s), speaker’s origin (see table 1), and as can be seen at a later point, to palate shape. All speakers are academics. Table 1 provides an overview of the participating subjects.

2.1. Measuring palatal morphology

The artificial palate of each subject was used to measure several morphological parameters. The Reading EPG system allows morphological data to be obtained since the electrodes in the artificial palate are placed with respect to anatomical landmarks (for more details see, Hardcastle, Gibbon, and Jones 1991). To do this, the palate was fitted to its dental cast and a high quality 1:1 photocopy was made for each speaker. The copy served as a reference to measure the horizontal (x) and vertical (y) coordinates of each of the 62 electrodes by means of a caliper.

The anterior palatal width (Width_ant) was defined as the horizontal distance between the two most peripheral electrodes in the first (anterior) row. The posterior palatal width (Width_post) was defined as the horizontal distance between the two most peripheral electrodes in the last row. The length of the palate (Length) was calculated as the vertical distance between the two most peripheral electrodes on the left side of the palate plus the vertical distance between those at the right side divided by two.
2.2. Acoustic and articulatory parameters

Based on the acoustic data we defined the beginning and end of the high frequency noise interval of intervocalic /s/ and calculated the acoustic fricative midpoint. Then we plotted the articulatory data across repetitions for each speaker at the fricative midpoint as contact frequency plots. In such a plot the percentage of tongue-palate contact of each electrode corresponds to a colour. White fields correspond to 0–25 percent contact, fields in light grey to 26–50 percent, fields in dark grey to 51–75 percent, and fields in black to 76–100 percent contact. Contact frequency plots for all speakers are shown in figures 4 and 5 in the results section.

The contact frequency plots and the morphological measures of the electrodes’ x- and y-coordinates served as the basis for calculating the constriction width (Constr) and provided an equivalent parameter for the
length of the front cavity (FrontCav). For measuring Constr, the electrode row which showed the narrowest air channel and the highest percentage of tongue-palate contact on both sides of the channel in the contact frequency plot was selected. For example, in EM3’s data (see figure 2, left) the row with the narrowest air channel is the first row, since only one electrode shows no contact (marked with arrows). In the second row a wider channel with two electrodes can be found. In cases where two rows had similar properties we chose the most anterior one. For instance in GM1’s data the third and fourth rows (see figure 2, right) both show three electrodes with no contact, but the third one was taken for further calculations. We then calculated the distance between the electrodes on both sides of the smallest constriction. A crucial factor for the reliability of this measure is the density of the electrode placement in the artificial palate. As can be seen in figure 1, the more anterior the electrodes, the closer they are to each other (on average for all speakers 2.89 mm in the first two rows, 3.13 mm in the third and fourth rows), and the more posterior the electrodes, the greater the distance between them (more than 4 mm in the posterior rows). A larger electrode distance increases the potential error since the tongue may be in contact with the palate at places between two electrodes. Hence the more back the articulation, the rougher the Constr parameter.

![Figure 2. EPG frequency plots for EM3 (left) and GM1 (right). The white arrows mark the row chosen to define the Constr parameter. The distance between the electrodes on the right and left sides of the constriction was used to calculate the constriction width. The black vertical arrow on the right (GM1) connects the electrodes which were used for calculating the FrontCav parameter.](image)

A similar issue arises for the calculated parameter FrontCav, where we subtracted the vertical placement (in mm) of the electrode in the channel showing minimal constriction from the corresponding electrode in the first
row. If the minimal constriction occurred in the first row, the FrontCav was zero. The front cavity length is not necessarily zero in reality, since the first electrode row is approximately 2–3 mm away from the necks of the teeth. Hence the FrontCav parameter is again a rough estimate and is shorter than the real front cavity length.

We also calculated the articulatory Centre of Gravity index (COG_ar) for the acoustically defined fricative midpoint. The COG_ar is a weighted index in the front-back dimension with higher values corresponding to a more front place of articulation and lower values to a more back articulation (Hardcastle, Gibbon, and Nicolaidis 1991). The index is described by the formula (where R = row):

\[
\text{COG}_{\text{ar}} = 0.5 \ R_8 + 1.5 \ R_7 + 2.5 \ R_6 + 3.5 \ R_5 + 4.5 \ R_4 + 5.5 \ R_3 + 6.5 \ R_2 + 7.5 \ R_1
\]

As can be seen, the more front the articulation, the more weight is added, e.g., in the first row R1 it is 7.5 while in the third row R3 it is only 5.5.

The three articulatory parameters COG_ar, FrontCav, and Constr were selected since we predict that they affect the acoustic properties of sibilants and may differ between males and females. As reported earlier, females often realize higher frequencies with a larger intensity than males (Flipsen et al. 1999). For the females we assume an underlying articulatory strategy with a more front articulation, a shorter front cavity length and/or a narrower minimal constriction while for males we suppose a more posterior place of articulation with a longer front cavity, and/or a wider constriction. The parameters COG_ar and FrontCav were both used since COG_ar is a normalized value and does not take into account individual palate length while FrontCav is an absolute value.

Choosing among the potential acoustic parameters for describing inter-speaker variation of /s/ was a challenging task, since there is no generally accepted parameter in the literature, and some of the parameters studied so far are partially redundant. Although partially redundant, we decided to use the main spectral peak (Peak), the acoustic Centre of Gravity (COG_ac), and Skewness. The Peak parameter was selected because Hughes and Halle (1956) suppose an inverse relation between the length of the front cavity and the frequency of the most prominent peak. The Peak was defined at the fricative midpoint in the LPC-smoothed spectrum of the frication noise with 3 LPC coefficients for the data with a sampling rate of 16 kHz and 4 coefficients for the data with a sampling rate of 22 kHz. The peak detection
was limited to the range of 2–10 kHz for the English speakers and the German speakers GM2, GM4, GF3 and GF6 and to 2–8 kHz for the remaining German speakers. The lower limit is intended to avoid detection of voicing, if any should occur. This limit should not interfere with the detection of peaks related to place of articulation, since they are generally located above 4 kHz for alveolars.

Second, although it is similar to the Peak, we calculated the acoustic Centre of Gravity (COG_ac) as the first spectral moment or mean by the formula

\[ \text{COG}_\text{ac} = \frac{\sum (F_i \times I_i)}{\sum I_i}, \text{ for } i = 2 \text{ to } 8 \text{ kHz} \]

where \( I \) is the amplitude and \( F \) the frequency of the spectrum. The choice of the COG_ac parameter was inspired by Tabain (2001), who reports a high correlation between the articulatory and the acoustic COG (see also Newman et al. 1991). Again, the range below 2 kHz was excluded in order to avoid detection of low frequency prominence due to voicing and the upper boundary was set to 8 kHz in order to compare English and German data, since the limits of the spectral range influence the calculation of COG_ac values.

Third, skewness was calculated following Forrest et al. (1988) since Flipsen et al. (1999), Stuart-Smith (2007) and Munson et al. (2006) report gender and sex differences with respect to this parameter. The frequency range from 700 Hz to the main spectral peak of each token per speaker was considered for the analysis. A single DFT for each window (window length 12.5 ms) was calculated at the fricative midpoint. We did not choose a larger fricative portion in order to compare articulation and acoustics at the same time point.

For statistical analyses we used SPSS (version 15.0). The details of the statistical procedures are given in the results section.

3. Results

3.1. Are there differences in male and female palatal dimensions?

To answer this question a MANOVA (Wilks-Lambda) was carried out in SPSS with all palatal parameters as dependent variables and language and sex as independent factors. The results provide evidence for a significant difference between palatal parameters of English and German speakers (\( F = \)
7.05, p = 0.002), but neither an effect for sex (F = 1.26, p = 0.318) nor an interaction between language and sex was found. English speakers differ significantly from Germans in anterior palatal width and palatal length (Width ant: F = 11.54, p = 0.003, Length: F = 4.48, p = 0.047); they have on average a longer, but narrower palate. A summary of the findings is given in table 2 below.

The fact that we did not find significant sex differences may lead to the conclusion that if any further acoustic or articulatory male-female differences in /s/ realization are found, they are grounded in sociophonetic factors. However, although not significant, we still found a trend for males to have longer palates than females for the English speakers only. The exceptions are two males who have a short palate. This finding is particularly interesting, since we assume that speakers should compensate for their palatal morphology in order to reach the required acoustic goal. Thus, if /s/ realization is a marker of gender identity, males with a short palate may compensate for it by means of a more back articulation to increase the front cavity length or by means of a wider constriction to have lower frequencies in comparison to females. In contrast, females with a relatively long palate may compensate for it with a more front articulation to decrease the length of the front cavity or by using a narrower constriction in order to realize higher frequencies in comparison to males.

Table 2. Descriptive statistics for the morphological parameters of the palate: Min = Minimum, Max = Maximum, Std. dev. = Standard deviation.

<table>
<thead>
<tr>
<th>Group</th>
<th>Parameter</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English males</td>
<td>Width_ant</td>
<td>11.80</td>
<td>15.50</td>
<td>13.56</td>
<td>1.34</td>
</tr>
<tr>
<td></td>
<td>Width_post</td>
<td>27.90</td>
<td>39.80</td>
<td>33.70</td>
<td>4.32</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>38.80</td>
<td>49.45</td>
<td>45.21</td>
<td>4.87</td>
</tr>
<tr>
<td>English females</td>
<td>Width_ant</td>
<td>13.30</td>
<td>15.20</td>
<td>14.08</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Width_post</td>
<td>30.70</td>
<td>39.10</td>
<td>33.13</td>
<td>3.02</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>37.15</td>
<td>42.80</td>
<td>40.39</td>
<td>2.09</td>
</tr>
<tr>
<td>German males</td>
<td>Width_ant</td>
<td>14.10</td>
<td>19.70</td>
<td>16.17</td>
<td>2.19</td>
</tr>
<tr>
<td></td>
<td>Width_post</td>
<td>31.50</td>
<td>50.00</td>
<td>39.00</td>
<td>6.73</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>36.60</td>
<td>47.45</td>
<td>40.02</td>
<td>3.81</td>
</tr>
<tr>
<td>German females</td>
<td>Width_ant</td>
<td>14.00</td>
<td>17.10</td>
<td>15.53</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>Width_post</td>
<td>30.50</td>
<td>39.00</td>
<td>33.96</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>34.10</td>
<td>43.65</td>
<td>39.03</td>
<td>3.87</td>
</tr>
</tbody>
</table>
If only the length of the hard palate mattered (biological factor), no compensation would be expected in the productions of males with short palates and of females with long palates. In this case, palatal length should correlate with place of articulation and front cavity length – as can be seen for the English speakers in our sample.

3.2. Evidence from articulation

Figure 3 displays scatterplots for the English and German speakers with the length of the palate on the y-axis and COG_ar (weighted articulatory index in the front-back dimension) on the x-axis. For the English speakers a strong negative correlation was found ($R = -0.78$, $r^2 = 0.58$, $F = 13.99$, $p = 0.004$). Palatal length predicts about 60 percent of the variance in place of articulation. This influence even increases when we consider not only the COG_ar as a speaker-normalized index, but also the absolute value of the front cavity length, FrontCav ($R = 0.79$, $r^2 = 0.63$, $F = 16.77$, $p = 0.002$). The two males who have a short palate behave articulatorily like the corresponding females and do not compensate for their morphology, which supports a biological explanation of the question under investigation.

For the German speakers no relation between palatal length and COG_ar was found. Independent of palate length, females consistently realize a more anterior articulation in comparison to males. In at least three cases,
German males (GM1, GM4, GM6) with a short palate length similar to that of females produce either a relatively posterior articulation or a wider constriction, suggesting a compensatory process.

Contact frequency plots for all the speakers were calculated and are displayed in figures 4 and 5. After a first evaluation two general findings become evident: First, females produce /s/ more anteriorly than males independent of their language. Second, German speakers realize a wider constriction than English speakers. For instance, most female speakers produce the minimal constriction in the first row of the palate; the plots for German females, however, more often show two electrodes with no contact corresponding to the air channel, whereas for the English females it is often only one electrode. Table 3 displays the descriptive statistics.

To provide further evidence several ANOVAs were carried out with one of the articulatory parameters as the dependent variable and language and sex as independent factors. COG_ar and FrontCav clearly showed an influence of sex (COG_ar: F = 8.87, p = 0.007, FrontCav: F = 12.58, p = 0.002), and Constr differed significantly between the two languages (F = 13.47, p = 0.002). All results are significant below the p < 0.016 level after Bonferroni correction.

Table 3. Descriptive statistics for the articulatory parameters: Min = Minimum, Max = Maximum, Std. dev. = Standard deviation.

<table>
<thead>
<tr>
<th>Group</th>
<th>Parameter</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English males</td>
<td>COG_ar</td>
<td>3.83</td>
<td>5.08</td>
<td>4.38</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Constr</td>
<td>6.00</td>
<td>10.60</td>
<td>7.20</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td>FrontCav</td>
<td>0.00</td>
<td>6.50</td>
<td>4.10</td>
<td>2.47</td>
</tr>
<tr>
<td>English females</td>
<td>COG_ar</td>
<td>4.39</td>
<td>4.75</td>
<td>4.60</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Constr</td>
<td>4.70</td>
<td>8.40</td>
<td>6.05</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>FrontCav</td>
<td>0.00</td>
<td>3.00</td>
<td>0.50</td>
<td>1.22</td>
</tr>
<tr>
<td>German males</td>
<td>COG_ar</td>
<td>3.83</td>
<td>4.40</td>
<td>4.21</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Constr</td>
<td>8.20</td>
<td>15.70</td>
<td>11.13</td>
<td>3.09</td>
</tr>
<tr>
<td></td>
<td>FrontCav</td>
<td>0.00</td>
<td>5.80</td>
<td>2.68</td>
<td>2.32</td>
</tr>
<tr>
<td>German females</td>
<td>COG_ar</td>
<td>4.44</td>
<td>5.30</td>
<td>4.74</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Constr</td>
<td>6.10</td>
<td>12.10</td>
<td>8.79</td>
<td>2.35</td>
</tr>
<tr>
<td></td>
<td>FrontCav</td>
<td>0.00</td>
<td>3.70</td>
<td>0.62</td>
<td>1.51</td>
</tr>
</tbody>
</table>

Hence, there is evidence for articulatory differences regarding place of articulation for males and females.
Figure 4. Frequency plots for all English speakers with the first two rows for males and the last two rows for females. Black markers correspond to 76–100% tongue-palate contact with respect to all the subject’s repetitions, dark grey markers correspond to 51–75%, light grey to 26–50%, and white markers to 0–25%. Upper incisors are located above the first row.
Since we also noted a potential covariation of COG_ar and FrontCav with palatal length (see figure 3 for English speakers), it is uncertain whether the
male-female differences are biological or sociophonetic in nature. To rule out possible biological factors, we ran the ANOVAs again, but included each time one of the palatal parameters as a covariate. If findings are still significant, the articulatory differences between males and females cannot be explained by morphological differences. Including palatal length as a covariate resulted in a decrease in the distinctive power, but findings are clearly significant with respect to sex and language (sex effect for COG_ar: F = 4.9, p = 0.039, FrontCav: F = 7.36, p = 0.014; language effect for Constr: F = 10.86, p = 0.004). Including Width_ant had only a marginal influence. The most pronounced influence was found for the Width_back covariate since it caused sex-specific differences not only in COG_ar and FrontCav, but also in constriction width (Constr: F = 4.48, p = 0.048).

To summarize, male-female differences in the articulation of /s/ still hold even if morphological factors are ruled out.

### 3.3. Results for acoustic parameters

An overview of the acoustic results is given in table 4. Results for the English females consistently show an extraordinarily high main spectral peak, on average above 8 kHz. This is quite different from the data for English speaking males and from all the German data, which do not show a male-female distinction in the Peak parameter. We do not believe that these differences between English and German females are due to the lower sampling frequency of the acoustic data for the Germans for the following reasons: (a) There were two German females for whom we had the same sampling frequency as for the English speakers, but peak values were still lower in comparison to the English females; (b) EPG data provide evidence for a wider constriction (Constr parameter) in the German data, which should affect the main spectral peak in the direction found in the acoustics, i.e. lower the main spectral peak (Shadle 1991); (c) Peak values were not included in the statistics for those cases where the front slope was still rising or did not decrease after the peak. For GF4, 10 repetitions were discarded from the recordings and for GF1 seven, but mainly because their spectrum was relatively flat. For all other German females main peaks were consistently found.
Table 4. Descriptive statistics for the acoustic parameters: Min = Minimum, Max = Maximum, Std. dev. = Standard deviation.

<table>
<thead>
<tr>
<th>Group</th>
<th>Parameter</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English males</td>
<td>Peak</td>
<td>4713</td>
<td>8140</td>
<td>6284</td>
<td>1229</td>
</tr>
<tr>
<td></td>
<td>COG_ac</td>
<td>4757</td>
<td>6167</td>
<td>5632</td>
<td>644</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-0.47</td>
<td>0.06</td>
<td>-0.25</td>
<td>0.20</td>
</tr>
<tr>
<td>English females</td>
<td>Peak</td>
<td>7041</td>
<td>9017</td>
<td>8282</td>
<td>722</td>
</tr>
<tr>
<td></td>
<td>COG_ac</td>
<td>5722</td>
<td>6856</td>
<td>6412</td>
<td>428</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-1.23</td>
<td>-0.47</td>
<td>-0.86</td>
<td>0.32</td>
</tr>
<tr>
<td>German males</td>
<td>Peak</td>
<td>4848</td>
<td>6785</td>
<td>5721</td>
<td>814</td>
</tr>
<tr>
<td></td>
<td>COG_ac</td>
<td>4006</td>
<td>6345</td>
<td>5463</td>
<td>903</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-0.89</td>
<td>-0.02</td>
<td>-0.44</td>
<td>0.31</td>
</tr>
<tr>
<td>German females</td>
<td>Peak</td>
<td>4778</td>
<td>6556</td>
<td>5841</td>
<td>716</td>
</tr>
<tr>
<td></td>
<td>COG_ac</td>
<td>5358</td>
<td>6366</td>
<td>5859</td>
<td>328</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td>-1.38</td>
<td>-0.20</td>
<td>-0.57</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Several ANOVAs were also carried out for the acoustic data. Findings for the Peak parameter showed an interaction of the two main effects language*sex (F = 6.6, p = 0.018) with differences between male and female English speakers, but none for the German speakers. However, it failed to reach significance after Bonferroni correction (p < 0.016). For the parameter Skewness an effect of sex was found (Skewness: F = 7.24, p = 0.014), again with more negatively skewed data for the female spectra. COG_ac did not reach significance after Bonferroni correction.

In summary, only the acoustic Skewness parameter showed significant differences between males’ and females’ /s/ realization. All other parameters showed a strong trend in the expected direction, but did not reach significance after correction for multiple tests.

3.4. Linking articulation and acoustics

Spearman Rho correlations showed that for the English speakers the most important articulatory correlate of the acoustics is the FrontCav (correlation with Peak: R = -0.755, p = 0.004, with COG_ac: R = -0.839, p = 0.001, with Skewness: R = 0.767, p = 0.004). It is followed by the articulatory COG_ar correlating with the acoustic COG_ac (R = 0.630, p = 0.028). This finding clearly supports earlier work from Hughes and Halle (1956), who suggest an inverse relation between front cavity length and the spectral peak. It is also in line with modeling work from Fant (1960), who mentions
the importance of the length of the front cavity for the spectral characteristics of /s/.

For the German speakers no correlation at all was found, i.e. there is at least no simple linear relation between articulatory and acoustic parameters. But German speakers also differ significantly from the English speakers in palatal length, anterior palate width, and with respect to the produced constriction width.

4. Discussion and conclusion

In order to answer the question as to whether differences that have been observed for male versus female /s/ realization are biological or sociophonetic in nature, articulatory, acoustic and morphological data were gathered from 12 English and 12 German speakers with 6 males and 6 females for each language.

In a first step we compared the palatal size parameters for males and females, assuming that this is the relevant part of the vocal tract where /s/ is realized and where males and females may potentially differ. Significant results for morphological parameters were not found that correlated with the sex of the speaker, but significant results were found which correlated with the language of the speaker. We are not certain whether these morphological differences are due to our small sample or whether they are representative for the differences between the German and English population. English speakers had on average a longer palate and narrower anterior palate width than Germans. Although not significant, we found a trend for English males to have a longer palate than English females, with two exceptions (two males with relatively short palates). This was particularly intriguing since we supposed that if /s/ distinctions are biological in nature, males with a short palate should not compensate for their palatal length. In contrast, if there are gender differences, males should compensate for their palate length, for instance by means of a back articulation or a wide constriction in order to decrease the high spectral energy. Pooling all data from males and females together yielded a particularly high correlation between palatal length and the length of the front cavity as well as the length of the front cavity and the acoustics for the English speakers. These results support the biological explanation of male-female differences in /s/ realization and speak for the need to obtain
speakers’ morphological data instead of simply splitting data into male and female results.

However, we not only found support for the biological explanation of /s/ differences, but also for the sociophonetic explanation, i.e. there was a mixture of effects. Potential biological influences were factored out by means of calculating several ANOVAs, where each time one morphological parameter was included as a covariate. Even if the morphological parameters reduced the power of the distinction, males and females still showed significant differences with respect to the articulatory COG_ar and FrontCav. Consequently, females actively produce a more front place of articulation and a shorter front cavity than males. These articulatory differences had an impact on the measured acoustic parameters for the English speakers. We found that the length of the front cavity was the most important parameter correlating with the main spectral peak, with the COG_ac, and with the spectral skewness.

Such a correlation was, however, not found for the German speakers, who differed from the English speakers with respect to palate size and constriction width: Although German females in most cases realized a similar anterior articulation to the English females, their constriction width was significantly wider. This also held true for the males. We assume that the wider constriction yielded the generally lower acoustic frequencies found in the German data.

An additional factor, which could be responsible for the language-specific differences, may be that the English phoneme inventory includes the neighboring /θ/ to the /s/, but the German inventory does not. Jongman, Wayland and Wong (2000) report lower spectral means for English /θ/ (averaged over speakers and vowel contexts ca. 5100 Hz) in comparison to /s/, and Narayanan, Alwan and Haker (1995) found a greater constriction area for /θ/ than for /s/. The way in which German /s/ differs in our data from English /s/ is in the direction of exactly these characteristics of /θ/. We therefore suggest that English /s/ may be more constrained than German /s/ in order to avoid perceptual confusion with /θ/.

Taken together, the findings of this study provide evidence for a mixture of effects contributing to the male-female differences in /s/ production. Palatal size parameters did not differ with respect to sex, but a trend was found for the English males to have a longer palate than the English females. Hence, differences in palatal size may generally be negligible, but in some cases dependent on the recorded sample. If morphological size
parameters are ruled out, male-female differences still remain in the articulatory production and acoustic realization of /s/.

Acknowledgements

We would like to thank Jana Brunner for making part of the data available to us, Jim Scobie, Alan Wrench, and Mariam Hartinger for their help in doing the recordings at QMUC Edinburgh, and Jörg Dreyer for his technical expertise during the recordings at ZAS Berlin and for measuring the placement of the electrodes in the artificial palates. Thanks to Olessia Panzyga and Anke Busler for analyzing acoustic data, to Jonathan Harrington and Stefanie Jannedy for providing background information on sociophonetics, to Ralf Winkler for scripting, to Caterina Petrone for her expertise on statistics, to Laura Koenig and Bernd Pompino-Marschall for proofreading, and to Keith Johnson and Marija Tabain for their comments on an earlier version of this paper. This work was supported by a grant from the German Research Council (DFG) PPO 334/4-1 and a grant from the DFH to the PILIOS project. It is dedicated to Dieter Fuchs.

References

Brunner, Jana, Susanne Fuchs, and Pascal Perrier

Fant, Gunnar

Fitch, Tecumseh W., and Jay Giedd

Flipsen, Peter Jr., Lawrence Shriberg, Gary Weismer, Heather Karlsson, and Jane McSweeny

Forrest, Karen, Gary Weismer, Paul Milenkovic, and Ronald N. Dougall
Fuchs, Susanne, Bernd Pomponio-Marschall, and Pascal Perrier

Gordon, Matthew, Paul Barthmaier, and Kathy Sands

Hardcastle, William J., Fiona Gibbon, and W. Jones

Hardcastle, William J., Fiona Gibbon, and Katharina Nicolaidis

Heffernan, Kevin
2004 Evidence from HNR that /s/ is a social marker of gender. Toronto Working Papers in Linguistics 23(2): 71-84.

Hughes, George W., and Morris Halle

Jannedy, Stefanie, and Jennifer Hay (eds.)
2006 Modelling sociophonetic variation Special Issue of Journal of Phonetics.

Johnson, Keith

Jongman, Allard, Ratree Wayland, and Serena Wong

Munson, Benjamin, Elizabeth C. McDonald, Nancy L. DeBoe, and Aubrey R. White

Narayanan, Shrikanth, Abeer Alwan, and K. Haker

Newman, Rochelle S., Sheryl A. Clouse, and Jessica L. Burnham


