

The perception of high frequency sibilants in Hungarian male speech

Péter Rác

New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Christchurch, New Zealand
peter.racz@canterbury.ac.nz

András Shepác

Eötvös Loránd University, Budapest
schepacz.andras@budling.hu

Abstract: The aim of this paper is to report on an experiment designed to evaluate the perception of high frequency sibilant articulations in Hungarian male speech and to theorise on the results. The main findings of the experiment are that the Hungarian listeners rate high frequency sibilants with femininity. These findings suggest that there is at least some social awareness of sibilant frequency in Hungarian. What follows from this is, in turn, that the sociolinguistic salience of sibilants as a variable is not confined to dialects of English, where the phenomenon has been most thoroughly described and discussed.

Keywords: sociolinguistics, gender, phonetics, sibilants, matched-guise

This paper reports on the results of a perception experiment in Hungarian. The hypothesis we set out to test is that Hungarian listeners associate higher frequency realisations of voiceless sibilants with percepts of femininity in male speech. This hypothesis conforms with a general line of research in English which shows that higher frequency realisations of sibilants are associated with percepts of femininity, female speech, or gay male speech. In order to test our hypothesis, we ran an experimental design in which participants listened to two voice clips which only differed in sibilant realisation and then rated the voices on attitudinal scales, including **masculine–feminine**. Our results support the hypothesis that listeners associate higher frequency sibilants with femininity. They also indicate that this is true without respect to the gender of the listener or their sexual orientation.

The paper is structured as follows. Section 1 reviews the relevant literature and spells out the hypothesis, section 2 discusses the experimental design and procedure, section 3 overviews the results, and section 4 provides a brief discussion.

1. Variation in sibilant realisation

This section discusses the characteristics of sibilant consonants, sociolinguistic and sociophonetic research on sibilants and (percepts of) sex and sexual orientation in English, and the context of investigating sibilants in Hungarian in the light of English findings.

1.1. The phonetic characteristics of sibilant sounds

Sibilant fricatives (and affricates) are produced by creating a constriction with the tongue nearing the palate. The sibilant /s/ is articulated with the tip of the tongue (the apex) in a dental position, with the tongue tip touching the upper teeth, or an alveolar position, with the tongue tip touching the ridge directly behind the upper teeth. (The place of articulation in most English dialects is alveolar.) A jet of air is blown through the constriction created by the tongue. This jet of air is directed to hit the teeth, thereby creating a noise source. The turbulence created this way is the principal source of this sound. The main characteristics of the sibilant /ʃ/ are similar, with the important difference that the place of articulation is posterior relative to /s/, typically postalveolar (Ladefoged & Maddieson 1996; Shadle 1991).

The acoustic image of a sibilant sound is a bulk of white noise centered around a given frequency range. The frequency-intensity spectrum of this can be described using a number of measures which we discuss in detail in section 2. The main point is that the frequency range is consistently different for /s/ and /ʃ/ and for female and male voices. For instance, in the study of Munson et al. (2006), the mean centre of gravity values of the energy spectrum of /s/ and /ʃ/ for female speakers are around 4200 Hz and 7500 Hz, respectively. For male speakers, these values are 3600 Hz and 6500 Hz. That is, /s/ has a higher frequency spectrum than /ʃ/, and female speakers have higher frequency spectra for both sounds.

The main determining factor in the realisation of /s/ is the position and shape of the tongue. Generally, a more laminal articulation leads to a higher frequency range of the noise, as well as a difference in its skewness

(the off-slant of the energy spectrum). Even though this shift from apical to laminal articulation does not entail the physical fronting of the tongue, it is often referred to as *fronting* in the literature (cf. Pharao et al. in press). It has to be noted that the relationship between the acoustic image of the sibilant and the position and shape of the articulator is not clear. For instance, apical articulation leads to lower centre of gravity values in English, but to higher ones in French (Dart 1991). Since we will focus on the acoustics in this paper, this problem is not directly relevant here.

Hungarian sibilants conform to the world trend (Siptár 1994): /s/ is dental, with a laminal tongue shape, while /ʃ/ is postalveolar, with slight lip-rounding. Hungarian is, in general, not very different from English in its sound structure. According to Siptár & Törkenczy (2000), the language has a pair of labial, dental, and palato-alveolar(/postalveolar) fricatives (with a voicing distinction) with the addition of velar /h/, also regarded as a fricative. The set of fricatives includes, beyond these contrastive segments, palatal fricative realisations of the glide /j/. One chief difference would be, as we will point out below, that Hungarian /s/ is dental, with the result that both sibilants are realised anterior relative to English.

1.2. Sociophonetic research on sibilant sounds

Higher peak frequency, higher centre of gravity and larger skewness of the energy spectrum, are, along with segmental duration, characteristics of /s/ in English that have been shown to correlate with (perceived) speaker sex and (perceived) speaker sexual orientation. Linville (1998) and Smyth et al. (2003) find a link between perceived sexual orientation and /s/ peak frequency: male voices with realisations that have a higher peak frequency are perceived as more gay-sounding. Munson et al. (2006) find an effect of the skewness of the energy spectrum of /s/ on percepts of sexuality. Mack and Munson (2012) also confirm that a higher frequency, more compact energy spectrum is rated as sounding more gay in male speech. The latter two papers show that higher peak frequency, centre of gravity, and a more negative skewness of /s/ is a characteristic difference between female and male speech: characteristic female realisations of /s/ are then similar to male realisations that are rated to be more feminine or gay. The only reference to this effect in any other language, to the authors' knowledge, is the ongoing research on Danish by Pharao et al. (in press).

The literature is extensive, and this mirrors the subject's complexity. While it seems to be the case that certain spectral characteristics correlate with both speaker sex and perceived speaker orientation, it is certainly not

the case that there would be an across-the-board gay male speech variant that heavily borrows from female speech. Most papers focus on /s/, but since the acoustic cues of /s/ and /ʃ/ are similar, hypotheses formed on the basis of the former can be straightforwardly tested on the latter. A general review of relevant results is provided by Smyth & Rogers (2008) and Mack & Munson (2012), while Campbell-Kibler (2011) expands on the problem of defining language variants using sets of individual variables, using the example of /s/.

Munson and Babel (2007) note that research on projecting sexual orientation through speech in languages other than English would be a welcome addition to existing sociophonetic work. If one wants to answer their call, however, one is immediately confronted with a number of serious issues. This is certainly true for Hungarian, and we assume that researchers in other, less-studied languages face the same challenges. The main problem in Hungarian is the absence of earlier, quantitative sociophonetic research on sexual orientation and speech.

The first part of our problem is the choice of the speech variable. We do not know whether any correlations between (for instance) a variable and perceived femininity are universal since they have only been observed in English so far. The reasons we picked sibilant realisation were twofold. First, the existence of work on Danish by Phrao et al. (in press) – combined with anecdotal evidence from other European languages and the authors' impressions about Hungarian – suggests that the social salience of sibilant realisation as a variable connected to projecting sexual orientation is not exclusive to English. Second, higher frequency sibilant realisations are socially salient in Hungarian, albeit regarded as a speech impediment, with unclear social connotations. If a variable in Hungarian is very similar to a variable in English, we may ask whether the Hungarian variable would have the characteristics that the English variable does seem to have.

The second part of the problem is what to ask. We have no evidence that Hungarians would react to labels such as *gay-straight* in a way comparable to results in English, both because the perception of these labels is culturally specific and because deviations from the heterosexual norm are less accepted in Hungary. We decided to opt for the labels *feminine-masculine* instead. We do not claim that these labels trivially map to *gay-straight* or even that they are comparable to a large degree (though Levon 2006 finds a strong correlation between listener judgements on the former and on the latter in an English context). We only argue that testing the correlation between higher frequency realisations of voiceless sibilants with percepts of femininity in male speech is a good exploratory step in an area of research that can be and should be expanded in the future.

2. Design and procedure

The perception test presented here is based on the matched-guise paradigm (Lambert et al. 1960), and is heavily influenced by Levon (2006). The matched-guise test is typically used to compare languages or dialects with each other. The reason we opted for it is that it allows us to test directly whether higher frequency sibilant realisations have feminine percepts in Hungarian male speech. It is true that our design does not allow us to tell whether higher frequency sibilant realisations are registered otherwise, such as speech defects, but since that claim is not central to this paper, we find that issue relatively unproblematic.

2.1. Stimulus design

Our matched-guise experiment consists of two kinds of sound clips: two guises and five fillers. Each clip is a Hungarian male voice reading the same text, a small, neutrally-toned composition about the benefits of pet ownership. Each clip is about 30 seconds long. The **guises** were read by the same person, while each fillers by different people, all male native speakers of Hungarian between the age of 20 and 30. The person reading the guises, a trained phonetician and a native speaker of Hungarian, read the **normal** guise in his normal speaking voice and the **matched** guise in a modified voice.

The speaker was instructed to read the matched guise with more laminal voiceless sibilant articulations, which have a higher frequency acoustic impression. The recordings were made with a head-mounted Sennheiser M3-EW microphone with Alesis io|2 pre-amplifier in a quiet room. In order to demonstrate that the main difference between the two guises is in the acoustic spectra of the voiceless sibilants, we provide a table of the acoustic differences between the two guises. We measured differences that have a potential effect on perceived sexual orientation or perceived masculinity/femininity, following Munson et al. (2006) and Smyth & Rogers (2008).

Specifically, we looked at values of the first two formants and duration of the cardinal vowels /i/, $n = 19$ (only 13 are included here, as 5 were deleted in the normal and 3 in the strident guise), /u/, $n = 5$, and /a:/, $n = 15$ (all per guise). We also looked at mean pitch and pitch variation measured on tokens of /a:/, along with the duration, centre of gravity (cog), skewness, kurtosis, and intensity of dental /s/, $n = 7$, and postalveolar /ʃ/, $n = 15$. Centre of gravity refers to the locus of the acoustic energy spectrum

of the fricative. Skewness is the off-slant of the fricative energy spectrum, and kurtosis is the sharpness of its peak. (For details, cf. Forrest et al. 1988.)

Since the speaker read the same text twice, differences are presented using pairwise t-tests. Some values are not paired in the recordings because segmental reduction rendered measuring them impossible. These values were excluded from testing. We used Praat (Boersma & Weenink 2005) for acoustic measurements and R for statistical analysis. T-values and degrees of significance are reported for significant differences in Table 1.

Table 1: Acoustic differences between the normal guise and the matched guise

Variable	T-value	Sign.
/i/:F1	2.62	*
/i/:F2	3.31	*
/ʃ/:cog	6.95	***
/ʃ/:skew	-4.67	***
/ʃ/:kurt	-5.7	***

If we exclude tokens of /i/ that are adjacent to a voiceless sibilant ($n = 5$), the difference in F1 is no longer significant, and the difference in F2 is smaller as well ($t = 2.62$; $p < 0.05$). We hypothesise that the difference in /i/ realisation is generally due to coarticulation effects in the matched guise, irrespective of whether a neighbouring sibilant is present. Centre of gravity is significantly higher, while skewness and kurtosis are significantly lower for /ʃ/ in the matched guise. One likely reason for the absence of significant differences for /s/ is the low token count of the segment in the guises. All the above listed spectral differences remain significant if we consider all voiceless sibilants together. While intensity is significantly different for both /s/ and /ʃ/, we do not report these values here as we think intensity measurements made on the recordings are unreliable. None of the other potentially relevant differences (as listed above) were significant across the two guises.

While we cannot exclude the possibility of further differences between the normal guise and the matched guise, these numbers show that the difference in the acoustic envelope of the voiceless sibilants (specifically, /ʃ/) is the most salient one. This can be confirmed by looking at Figure 1. This figure shows the energy spectra of all /ʃ/ tokens in the normal guise

(straight line) and the matched guise (dotted line) after LPC smoothing with 30 peaks in Praat. We can see that the main energy of [ʃ] tokens in the matched guise is distributed in a higher frequency range. This comes with the acoustic impression of a higher frequency of the noise which is the main acoustic cue of sibilants.

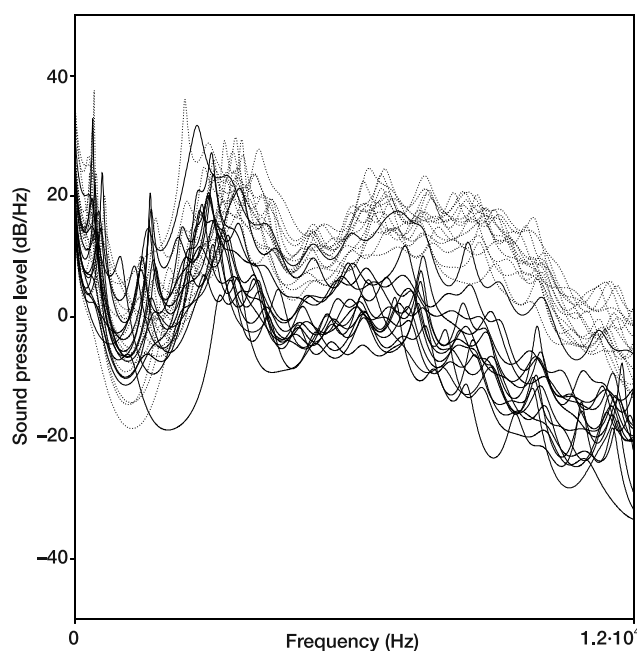


Figure 1: Energy spectra of /ʃ/ tokens in the normal guise (straight line) and the matched guise (dotted line)

We opted for the use of a trained speaker instead of other ways of manipulation to create the stimuli because we considered the other available options (such as splicing in modified tokens) too artificial for the given task. Trained speakers have been successfully used in similar earlier work (cf. Purnell et al. 1999) and we believe to have shown that, in our case, the acoustically most salient difference is the desired one – the sound spectra of voiceless sibilants. In a way, the above discussion presents the difference between the two guises in a clearer way – being explicit about acoustic detail – than what would be expected from a design using acoustic manipulation.

2.2. Task and participants

In our design, participants enter a room individually and listen to the seven unique voice clips wearing full-size headphones. The instructions imply that the participants are listening to seven different voices. The order of clips is set but each participant can listen to a clip multiple times. Half the total number of participants listened to the normal guise as second and the matched guise as seventh. The other half had the matched guise as second and the normal guise as seventh.

The participants rate the voices along eight attitudinal scales. The scales, shown in (1), are presented as horizontal lines on a paper sheet, in this set order.

- (1) generous–selfish
prude–promiscuous
feminine–masculine
friendly–unfriendly
orderly–messy
nice–mean
caring–insensitive
genuine–fake

The scales and their ordering were largely based on Levon (2006), but we excluded his gay–straight dimension for reasons discussed in section 1. We only have predictions about two of the scales. We predict that the matched guise would be rated as less masculine. We also predict that the matched guise is not rated more ‘fake’. Comparatively higher ratings on ‘fake’ can be interpreted in two ways. First, they might imply that the listener found the matched guise unnatural or forced. Second, they might imply that the listener has no assumptions about the naturalness of the recording but finds the projected personality behind the matched guise to be dishonest or simply too mannered. This interpretation problem, however, only arises if there is a significant difference between the guises on this scale. An absence of such a difference, we believe, implies that participants who did not notice that they heard the same person twice did make an implicit distinction either.

After the experiment each participant was de-briefed. Those who realised they heard the same person twice were excluded from analysis. The experiment had 34 participants, 4 of which were excluded (3 had the normal guise as second and 1 the matched guise as second in the sequence). The remaining 30 were 10 straight women, 10 straight men, and 10 gay men, between the age of 20 and 30, from the capitol city, sexual orientation self-identified.

3. Results

Judgement ratings were standardised for participants across all judgements. Figure 2 shows the standardised ratings for all attribute scales for the two guises (normal and matched).

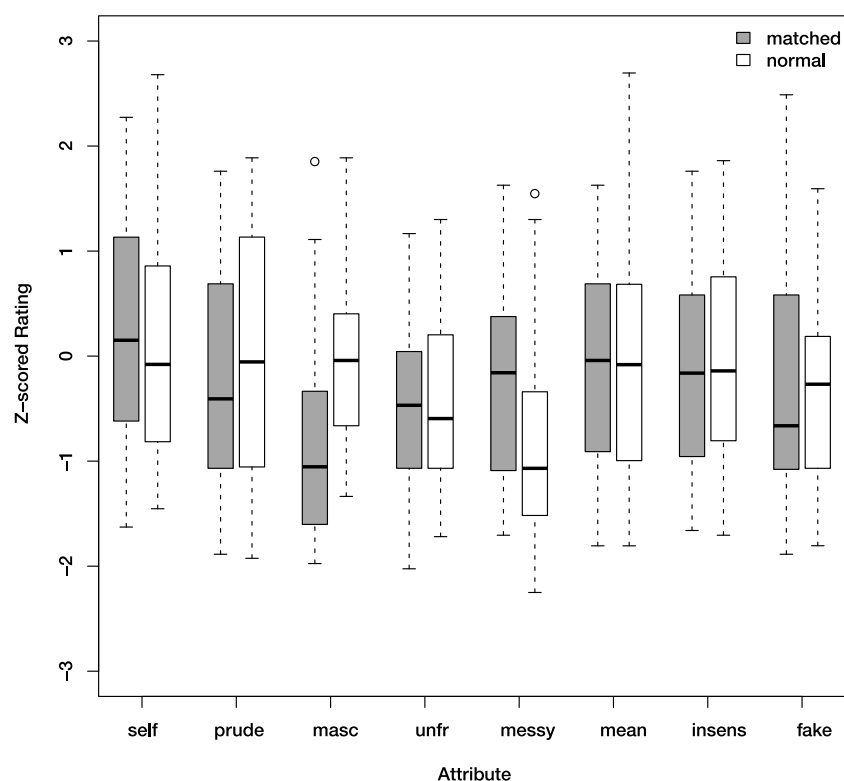


Figure 2: Standardised ratings for all participants ($n = 30$)

We have fitted analysis of variance models with repeated measures (with **condition** nested in **subject**) on the standardised data. The dependent variable is rating **score** on each of the scales, the factors are **condition** (type of guise), **group** (whether the listener is a straight woman, a straight man, or a gay man), and the interaction of condition and group. Two negative results should be noted here. First, the matched guise was not rated more fake than the normal guise, which suggests that those participants who did not realise that they had heard the same person twice did not do so

even implicitly. This was true irrespective of whether the matched guise came second or last in the experimental sequence.

Second, we expected that participant sexual orientation would have an effect on the ratings. This did not turn out to be the case – the interaction of group and condition was not significant for any of the scales. (We observed that straight women tended to rate the guises differently from the rest of the participants on the ‘promiscuous–prude’ dimension. It has been pointed out to us, however, that this scale is not comparable to the other scales in the sense that the two end points are not polar opposites of each other in the same way as in the case of e.g., ‘feminine–masculine’, so we exclude it from this discussion.)

Table 2 sums up the F values and degrees of significance for all participants across paired differences in all ratings in the two conditions. Both significant and non-significant values are reported.

Table 2: Differences in ratings between guises for all participants

Scale	F -value	Sign.
generous–selfish	0.294	n.a.
prude–promiscuous	0.009	n.a.
feminine–masculine	14.308	***
friendly–unfriendly	0.83	n.a.
orderly–messy	9.084	**
nice–mean	0.224	n.a.
caring–insensitive	0.579	n.a.
genuine–fake	0.826	n.a.

The two significant results are that the participants rated the matched guise lower on the ‘feminine–masculine’ and higher on the ‘orderly–messy’ scales. The direction of the effect can be seen in Figure 2.

The results support our main hypothesis that higher frequency sibilant articulations are regarded as more **feminine** in Hungarian male speech. The result that they are also seen as more **messy** is surprising. Since we could – at best – come up with a post hoc hypothesis about why this is so, we take note of this result but do not discuss it further.

4. Discussion

We believe the results reported here should be interpreted narrowly, as evidence for a link between high frequency voiceless sibilant spectra and the percept of femininity in Hungarian male speech. It is clear, however, that there is a far larger potential for this line of study. Results can be more reliable if we better control for the difference between the guise and the matched guise, build in complete randomisation without repetition for order of guises and order of scales, and include more participants in the study. In parallel with the body of research outlined in section 1, it is also possible to formulate a series of research questions, such as on the interpretation of the 'gay–straight' dimension and its relationship to femininity, the potential of other variables to carry connotations of femininity in Hungarian, or whether, in general, sexual orientation is projected through speech and what shapes this projection might assume.

Further research in this area is not only useful in shedding more light on the connections between speech, gender, and sexuality in Hungarian, but it also allows us to see whether the findings of quantitative sociophonetic work in English carry over to other languages and can thus claim universality in some sense. We hope to have made a first step in this direction.

Acknowledgements

The project benefitted immensely from the criticism of Patrick LaShell, Viktória Papp, and Amanda Cardoso. We are also grateful for the feedback we received at workshops in Berlin, Edinburgh, and Christchurch, as well as for the comments of our reviewers. All faults remain ours.

References

- Boersma, Paul and David Weenink. 2005. Praat: Doing phonetics by computer. [Computer program]. <http://www.praat.org/>.
- Campbell-Kibler, Kathryn. 2011. Intersecting variables and perceived sexual orientation in men. *American Speech* 86. 52–68.
- Dart, Sarah N. 1991. Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics* 79. 71–94.
- Forrest, Karen, Gary Weismer, Paul Milenkovic and Ronald N. Dougall. 1988. Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America* 84. 115–123.

- Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. Cambridge MA & Oxford: Blackwell.
- Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner and Samuel Fillenbaum. 1960. Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology* 60. 44.
- Levon, Erez. 2006. Hearing 'gay': Prosody, interpretation, and the affective judgments of men's speech. *American Speech* 81. 56–78.
- Linville, Sue Ellen. 1998. Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia Phoniatica et Logopaedica* 50. 35–48.
- Mack, Sara and Benjamin Munson. 2012. The influence of /s/ quality on ratings of men's sexual orientation: Explicit and implicit measures of the 'gay lisp' stereotype. *Journal of Phonetics* 40. 198–212.
- Munson, Benjamin and Molly Babel. 2007. Loose lips and silver tongues, or, projecting sexual orientation through speech. *Language and Linguistics Compass* 1. 416–449.
- Munson, Benjamin, Elizabeth C. McDonald, Nancy L. DeBoe and Aubrey R. White. 2006. The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics* 34. 202–240.
- Pharao, Nicolai, Marie Maegaard, Janus Møller and Tore Kristiansen. in press. Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43.
- Purnell, Thomas, William Idsardi and John Baugh. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18. 10–30.
- R Development Core Team. 2009. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. <http://www.R-project.org>.
- Shadle, Christine H. 1991. The effect of geometry on source mechanisms of fricative consonants. *Journal of Phonetics* 19. 409–424.
- Siptár, Péter. 1994. A mássalhangzók [Consonants]. In F. Kiefer (ed.) *Strukturális magyar nyelvtan 2: Fonológia [A structural grammar of Hungarian 2: Phonology]*. Budapest: Akadémiai Kiadó. 183–272.
- Siptár, Péter and Miklós Törkenczy. 2000. *The phonology of Hungarian*. Oxford: Oxford University Press.
- Smyth, Ron, Greg Jacobs and Henry Rogers. 2003. Male voices and perceived sexual orientation: An experimental and theoretical approach. *Language in Society* 32. 329–350.
- Smyth, Ron and Henry Rogers. 2008. Do gay-sounding men speak like women? *Toronto Working Papers in Linguistics* 27. 129–144.