The production and perception of the German /s, ç, ʃ/ contrast

Felicitas Kleber[1], Mary Lowery[2], Rebecca Stegmaier[1]

[1]Institute of Phonetics and Speech Processing, LMU München, [2]Case Western Reserve University

kleber@phonetik.uni-muenchen.de, mml103@case.edu, stegmaier@phonetik.uni-muenchen.de

Abstract

This study investigated the production and perception of the rare three-way /s, ç, ʃ/ contrast in Standard German which has been reduced to a two-way contrast in a number of German varieties. All three fricatives differed in their spectral properties with /ç/ taking up an intermediate position and featuring a higher F2-transition. In perception, listeners used spectral cues more categorically to distinguish between /s/ and /ʃ/ than they used the F2-transition for the /ç, ʃ/-distinction. Cues were weakened in perception when listeners saw pictures of a neutralizing dialect. Such cue weakening may be related to diachronic changes by which /ç/ merges with /ʃ/.

Introduction

Standard German features a three-way contrast between alveolar /s/, postalveolar /ʃ/, and palatal /ç/ with the latter being rare in the world's languages and even more so as a contrasting sound to /ʃ/ (Jannedy & Weirich, 2016). The /ç, ʃ/ opposition has merged to /ʃ/ or /ɕ/ in many Central German varieties (Hall, 2014), e.g. Hessian. This merger may display similarities to other mergers in sibilant contrasts as observed e.g. for Polish. Bukmaier and Harrington (2016) argue that the instability of Polish retroflex /ʂ/ might come about because it shares acoustic characteristics with the other two sibilants (low formant frequency transition with dental /s/ and lower center of gravity with alveolo-palatal /ɕ/), while /s/ and /ɕ/ have unequivocal acoustic characteristics (/s/: high center of gravity; /ɕ/: high formant frequency transition). Jannedy and Weirich (2016) acoustically analyzed all five German voiceless fricatives (i.e. the above mentioned plus /f/ and /χ/) in German speakers of a neutralizing (i.e. where /ç, ʃ/ have merged) and a non-neutralizing variety and found that speakers of the neutralizing variety realized the contrasts between a number of fricative pairs to a lesser extent than speakers from the non-neutralizing variety (though not necessarily /ç, ʃ/). Additionally, they noted that the Discrete Cosine Transformation (DCT, and here especially DCT2, i.e. the curvature of the spectrum) best captures the /s, ç, ʃ/ opposition (cf. also Bukmaier & Harrington, 2016), but that it remains difficult to quantify the /ç,ʃ/ contrast even in speakers of a non-neutralizing variety. That is, generally greater overlap in the fricative noise may make this contrast less salient. Apart from these, however, acoustic investigations of the German fricatives are sparse (e.g. Jannedy & Weirich, 2014), which is why this study aims at further contributing to a better understanding of (1) the phonetic implementation of this contrast in production by Standard German speakers and (2) the role as well as (3) the stability of these cues in the perception of Standard German listeners.

While Polish listeners have been found to use formant transitions in vowels adjacent to fricatives to distinguish between /ɕ/ and /ʂ/ (Nowak, 2005), German listeners are said to exploit vocalic cues to fricative contrasts to a much lesser extent (Lipski & Mathiak, 2007; Wagner, Ernestus & Cutler, 2006). To investigate this further and to achieve our second aim we therefore tested the perceptual role of both the energy distribution within the fricative noise and that of formant transitions that are likely to differ in the vicinity of palatal /ç/ vs. (post)alveolar /s, ʃ/.

Our third aim is motivated by sound change theories that see the source for diachronic change in the waning of certain perceptual cues (Beddor, 2009) and by the study of Jannedy and Weirich (2014), who found that older more than younger listeners from Berlin rated identical acoustic stimuli differently as a function of co-presented information that could be associated with speakers who either neutralize the contrast or not. They interpreted this finding as being indicative of a sound change in progress by which /ç/ and /ʃ/ merge. The present study partly replicates this earlier study but tests it with Standard German participants of the same generation who were primed with pictures of different varieties. Such priming studies allow for an indirect

simulation of language contact as a trigger for change.

Speech Production

Method

Acoustic data were collected from ten female native speakers of Standard German (aged 18 to 27, mean age = 21.6) using the SpeechRecorder Software (Draxler & Jänsch, 2004, version 3.4.21) with a 44.1 kHz sampling rate and a 16-bit resolution in a sound-attenuated booth at the Institute of Phonetics at LMU Munich. All participants were students of Phonetics who had been living in Munich, Germany, for at least six months at the time of the recording and who had no speaking or hearing deficits.

The participants were asked to produce the words *wissen* (/vɪsən/ 'to know'), *wichen* (/vɪçən/ 'to give way' 3PL PRETERITE), and *wischen* (/vɪʃən/ 'to mop') three times each along with 40 other filler words in randomized order. The words were presented one at a time on a computer screen and participants were asked to read the word out loud at a normal pace. A total of nine single-word utterances per speaker was available for analysis.

All recordings were first segmented automatically using WebMAUS (Kisler et al., 2016) and subsequently all segments were checked in Praat (version 6.0.16, Boersma & Weenink, 2016) and corrected according to standardized IPS criteria. The on- and offset of the vowel coincided with a clearly visible 2nd formant frequency (F2) and the on- and offset of the fricative (the fricative's onset being identical with the vowel's offset) were set to the beginning and end of aperiodicity in the signal. All data were then converted into an EMU format to allow for acoustic analyses in R using the emuR package (Winckelmann & Raess, 2014).

F2 at 80% of the preceding vowel and the 2nd (curvature) mel-scaled DCT coefficient (hereafter mel-DCT2; cf. the methodology in Bukmaier & Harrington, 2016) were the dependent variables in two separate repeated measures ANOVAs with fricative (/s/ vs. /ç/ vs. /ʃ/) as the fixed factor and speaker as the random factor.

Results

Commensurate with Figure 1 Standard German speakers distinguish all three fricatives by means of mel-DCT2 ($F[1,9]=83.3$; $p < .001$) with /ç/ taking up an intermediate position between /s/ and /ʃ/.

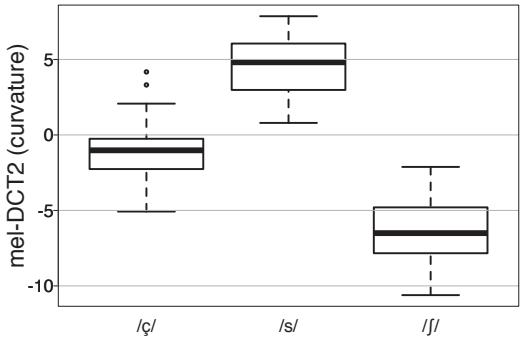

Figure 1. Mel-DCT2 separately for palatal /ç/*, alveolar* /s/*, and postalveolar* /ʃ/*.*

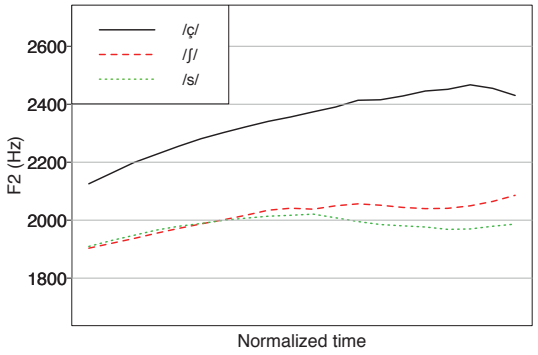

Figure 2. F2 trajectory over the entire /ɪ/ *vowel separately for alveolar* /s/ *(green, dotted), palatal* /ç/ *(black, solid), and postalveolar* /ʃ/ *(red, dashed).*

Figure 2 shows that F2 is another important cue that sets /ç/ apart from the other two fricatives. From the very beginning and throughout the entire vowel, F2 is considerably higher in the former fricative compared to the latter two which, in turn, do not differ in that parameter. In addition to a significant effect for fricative on F2 ($F[1,9]=137.8$), post-hoc tests revealed significant differences between /ç/ and /ʃ/ as well as /ç/ and /s/ but not between /s/ and /ʃ/. The next section thus asks among other things whether listeners use this cue to separate the palatal from the postalveolar fricative.

Speech Perception

Method

A different set of 26 speakers (aged 20 to 28, mean age = 23.8; five male) of Standard German, of which the majority came from the Munich area and two from the state of North-Rhine Westphalia, participated in the

perception experiment. They rated stimuli (six repetitions each) from two 13-step continua that were generated using the morphing algorithm provided by STRAIGHT: the first continuum, ranging from *wisst* (/vɪst/ 'know' 2PL[1], stimulus 1) to *wischt* (/vɪʃt/ 'sb. mops' 3SG/2PL, stimulus 13) and commensurate with Figure 1 possibly encompassing *Wicht* (/vɪçt/ 'gnome'), was created by morphing the entire fricative between two selected natural productions of *wisst* and *wischt* by a male speaker. That is, not only mel-DCT2 but also other parameters present in the fricative were morphed. More importantly, however, only the fricative noise varied in the continuum while the rest of the word corresponded to one of the two selected productions and featured an ambiguous F2 in the preceding vowel. Subsequent analyses of the stimuli in emuR revealed that the continuum covered and slightly extended the range of mel-DCT2 values shown in Figure 1. Stimuli from this continuum were judged in a three-alternative-forced-choice (3AFC) task, where listeners had to choose one of the three words *wisst*, *Wicht*, and *wischt*.

Figure 3. Hessian (top row from left to right: Maddin Schneider, Frankfurt green sauce, city hall of Frankfurt) and Northern German (bottom row from left to right: Otto Walkes, lobscouse, Hamburg fish market) primes.

For the second continuum a high rising vowel-consonant (VC) F2-transition in *Wicht* was morphed into a low falling VC F2-transition in *wischt* (stimulus 13), such that only the transition phase was manipulated and all other parts of the stimulus were kept equal. This time the stimulus contained a fricative where mel-DCT2 was intermediate between typical /ʃ, ç/ values (cf. Figure 1). Subsequent measurements of F2 at the offset of the vowel showed that the F2 continuum ranged from F2 values of 2850 Hz (Stimulus 1) and 2490 Hz (Stimulus 13), respectively, i.e. the step size was 30 Hz. Stimuli from this continuum were judged in a 2AFC task, where listeners had to choose between *Wicht* and *wischt.*

Every listener was assigned to one of two groups each consisting of 13 listeners. Upon presentation of each auditory stimulus listeners saw pictures evoking an association with either a neutralizing dialect (here Hessian) or a non-neutralizing dialect (here Hamburg = Northern German) and were asked to judge which of the three or two orthographically presented words, respectively, they saw alongside with the pictures they had perceived. Three pictures per variety were used: both sets contained pictures of famous speakers, dishes, and characteristic sights typical for the respective region (cf. Figure 3). To maintain listeners' attention throughout the experiment each task was interspersed with six questions relating to the pictures and six questions relating to the words, respectively, each time presented on the previous slide.

Results

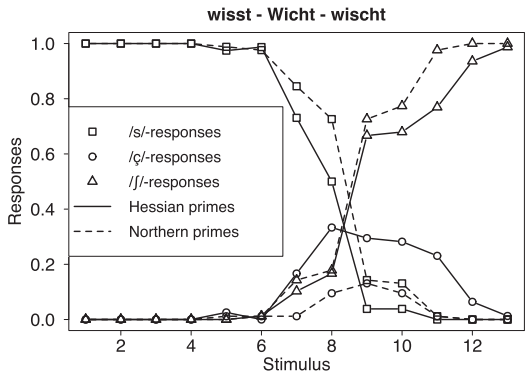

Figure 4. Response curves to the spectral continuum from wisst *to* wischt *separately for the two groups presented with Hessian (solid line) and Northern German primes (dashed line), respectively.*

Listeners used the spectral cues categorically to identify /s/ and /ʃ/ but not /ç/ (cf. Figure 4); listeners picked the latter category a few times when listening to stimuli 7 to 11, which were apparently perceptually ambiguous between /s/ and /ʃ/, but given the overall low percentage of /ç/ responses these stimuli were far from representing good /ç/ exemplars. The response curves for /s/ and /ʃ/ were very similar with respect to slope and category boundary, the only difference arising

[1] 2pl forms were used because the same experiment was run as a picture naming task with preschool children to whom the noun *Wicht* /vɪçt/ 'gnome' tested in the second continuum is more familiar than *wichen*.

from more /ç/-responses in the group presented with Hessian primes. This finding suggests that this group was less confident in giving /ç, ʃ/-responses.

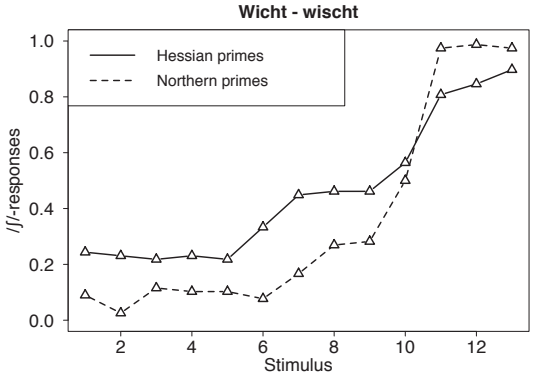

Figure 5. Response curves to the vocalic continuum from Wicht *to* wischt *separately for the two groups presented with Hessian (solid line) and Northern German primes (dashed line), respectively.*

Judgments of the vocalic continuum, on the other hand, turned out less categorical, in particular with respect to the middle part of the continuum (cf. Figure 5). The endpoints, however, were predominantly rated as *Wicht* (Stimulus 1) and *wischt* (Stimulus 8), respectively, i.e. the vocalic cue F2 indeed enables listeners to distinguish between /ç/ and /ʃ/. In this continuum, priming affected listeners' responses such that Northern German primes triggered a better perceptual separation than Hessian primes.

Discussion and Conclusion

Three main findings arise from this study: (1) Standard German speakers distinguish /s, ç, ʃ/ by means of mel-DCT2 whereby /ç/ takes up an intermediate position but is, in turn, additionally cued by a significantly higher F2. (2) Listeners use this vocalic cue to a greater extent than the spectral cue to identify /ç/. (3) Perception was in general less categorical in response to the F2 continuum compared to the spectral continuum and even more so when listeners saw Hessian primes. No such obvious priming effect was found for the /s, ʃ/ distinction along the spectral continuum except for a tendency towards greater /ç, ʃ/ confusion.

The first two findings clearly show that vocalic cues play a role in the identification of the rare palatal fricative and that listeners make use of this vocalic cue, if only in tandem with secondary cues. Our third finding demonstrates that listeners are affected by their (implicit) knowledge of dialectal differences and it adds to the existing evidence that perceptual boundaries are flexible (Jannedy & Weirich, 2014). Most importantly, the ability of listeners of a non-neutralizing variety to differentiate the three-way contrast is weakened in the presence of visual primes associated with contrast neutralization. Such perceptual cue weakening in a contact situation may be related to diachronic changes by which /ç/ merges with /ʃ/.

References

Beddor, P. S. (2009). A coarticulatory path to sound change. *Language, 85(4)*, 785–821.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.16, retrieved 12 April 2016 from http://www.praat.org/

Bukmaier, V., & Harrington, J. (2016). The articulatory and acoustic characteristics of Polish sibilants and their consequences for diachronic change. *Journal of the International Phonetic Association*, *46*, 311–329.

Draxler, C., & Jänsch, K. (2004). Speechrecorder – a universal platform independent multi-channel audio recording software. *Proc. LREC'04.*

Hall, T. A. (2014). Alveolopalatalization in Central German as markedness reduction. *Transactions of the Philological Society, 112(2),* 143–166.

Jannedy, S., & Weirich, M. (2014). Sound Change in an urban setting: Category Instability of the palatal fricative in Berlin. *Journal of Laboratory Phonology*, *5*, 91–122.

Jannedy, S., & Weirich, M. (2016). The acoustics of fricative contrasts in two German dialects. *Proc. P&P12*.

Kisler, T., Reichel U. D., Schiel, F., Draxler, Ch., Jackl, B., & Pörner, N. (2016). BAS Speech Science Web Services - an Update of Current Developments. *Proc. LREC'16.*

Lipski, C., & Mathiak, K. (2007). A magnetoencephalographic study on auditory processing of native and nonnative fricative contrasts in Polish and German listeners. *Neuroscience Letters, 415,* 90–95.

Nowak, P. M. (2005). The role of vowel transitions and frication noise in the perception of Polish sibilants. *Journal of Phonetics, 34,* 139–152.

Wagner, A., Ernestus, M., & Cutler, A. (2006). Formant transitions in fricative identification: The role of native fricative inventory. *J. Acoust. Soc. Am., 120(4)*, 2267–2277.

Winkelmann, R., & Raess, G. (2014). Introducing a Web Application for Labeling, Visualizing Speech and Correcting Derived Speech Signals. *Proc. LREC'14.*