Title: Do talkers produce less dispersed phoneme categories in a clear speaking style?

Outi Tuomainen
Dept of Speech, Hearing and Phonetic Sciences, University College London (UCL), Chandler House, 2 Wakefield Street, London WC1N 1PF, UK.
o.tuomainen@ucl.ac.uk

Valerie Hazan
Dept of Speech, Hearing and Phonetic Sciences, University College London (UCL), Chandler House, 2 Wakefield Street, London WC1N 1PF, UK.
v.hazan@ucl.ac.uk

Rachel Romeo
Speech & Hearing Bioscience and Technology, Division of Medical Sciences, Harvard University, TMEC 435, 260 Longwood Avenue, Boston MA 02115
rachelromeo@fas.harvard.edu

Running title: Phoneme category structure in clear speech

Abstract

This study investigated whether adaptations made in clear speaking styles result in more discriminable phonetic categories than in a casual style. Multiple iterations of keywords with word-initial /s/-/ʃ/ were obtained from 40 adults in casual and clear speech via picture description. For centroids, cross-category distance increased in clear speech but with no change in within-category dispersion and no effect on discriminability. However, talkers produced fewer tokens with centroids in the ambiguous region for the /s/-/ʃ/ distinction. These results suggest that, whereas interlocutor feedback regarding communicative success may promote greater segmental adaptations, it is not necessary for some adaptation to occur.

PACS number: 43.70.Mn

1. Introduction

When speakers are asked to speak clearly, either in an experimental setting or when communicating with interlocutors who have a hearing impairment or who are non-native, they make adaptations to various acoustic-phonetic characteristics of their speech. Typically, speech produced in an 'instructed' clear speaking style is slower, more intense, more hyperarticulated, with greater pitch variations and contains more frequent pauses than speech produced in a casual or conversational style [for a review, see Cooke et al., 2014]. One issue of interest is the degree to which these clear speech adaptations are aimed at making phonetic categories more easily discriminable rather than just resulting in global enhancements.

Greater distinctiveness between categories could be achieved in two ways [see also discussion by Newman et al., 2001]. The first is by increasing the distance between category distributions. For example, for a fricative /s/-/ʃ/ contrast, this may entail increasing the difference between the mean fricative center of gravity across distributions comprising multiple iterations of /s/ and /ʃ/. Some studies have used experimental designs where a clarification of a phonetic contrast was at least implicitly elicited by having talkers produce words in response to a miscomprehension by a real or simulated interlocutor. In such studies, there is consistent evidence that talkers do produce a greater distance between the two categories [e.g., Maniwa et al., 2009; Buz et al., 2016].

In studies where talkers are instructed to speak clearly, or carry out a task that requires them to adopt a clear speaking style, but where there is not a specific focus on phonetic contrasts, clear speech modifications have also been shown to include enhancements at the segmental level, in both vowels and consonants [e.g., Smiljanic & Bradlow, 2008; Granlund et al., 2011].

A second strategy to increase distinctiveness would be to be more consistent in phoneme production: This would reduce within-category dispersion and potential category overlap. In perception experiments, slower reaction times have been obtained for talkers who had greater within-category variability, thus supporting the relevance of consistency in production for speech perception [Newman et al., 2001]. In a recent study, when producing target words containing a voiceless plosive (e.g., *pill*) for the benefit of an interlocutor in the presence of potentially confusable foils containing a voiced plosive (e.g., *bill*), talkers reduced the number of potentially confusable tokens by reducing variance at the extremes of the category distribution, especially at short VOTs [Buz et al., 2016]. In their 'adaptive speaker framework', Buz et al. argued that the presence of interlocutor feedback was key in eliciting these adaptations as the extent to which these clarifications occurred was dependent on the communicative success of previous exchanges between the two talkers. As within-category dispersion has received relatively little attention, the current study investigated whether the production of a clear speaking style in the absence of an interlocutor also involved such a strategy, or whether, on the contrary, the effort to produce more distinct consonants led to increased variability in production. The /s/-/ʃ/ fricative place contrast was chosen for analysis. This phonetic contrast is primarily marked by spectral differences between the two consonants, unlike stop voicing contrasts such as /p/-/b/ where the primary cue, voice onset time, is a durational cue. With durational cues, there are concerns about potential contrast enhancement effects being strongly influenced by changes in articulation rates across conditions, as also noted in Granlund et al. (2013). Fricatives produced in casual and clear styles have been exhaustively analyzed in a study involving 500,000 measurements of fricative tokens [Maniwa et al., 2009] but although that study considered a wide range of acoustic cues to this contrast, it focused on cross-category distance and gave less attention to within-category dispersion.

Here, we focus on the changes in internal category structure in adult speech that result from adopting a clear speaking style. More specifically, the aim of this study was to establish whether speech production in a clear speaking style entails an increase in discriminability of phonetic contrasts, even in the absence of genuine communicative behaviour that emerges when there is feedback from the interlocutor. We also considered whether any increased discriminability was as a result of greater internal consistency in the way in which consonants are produced in a clear speaking style, thus producing less dispersed categories as well as increasing category distance. As women have been found in a number of studies to be more intelligible than men [Bradlow et al.,1996; Hazan & Markham, 2004], we also examined the data for sex effects to see whether women showed evidence of making greater segmental adaptations in clear speech than men. The LUCID corpus is a large corpus of spontaneous and read casual and clear speech [Hazan & Baker, 2011]. It includes on average 32 iterations per consonant for each of 40 adult talkers for the /s/-/ʃ/ fricative place contrast in word-initial position, obtained via picture elicitation; thus, it enables us to examine within-category dispersion and across-category distance in two 'instructed' speaking styles.

2. Method

2.1 Talkers

Talkers were forty adults (20 M, 20 F), with a Southern British English accent aged between 19 and 29 years old; they were all university students or faculty. They were screened for normal hearing thresholds and reported having no language impairment.

2.2 Materials

The LUCID corpus materials include 18 minimal or near minimal pairs; nine pairs contain the phonemes /s/-/ʃ/ in word-initial position while the remaining nine are pairs with /p/-/b/ in word-initial position. An easily-recognizable picture was found for each of the 36 keywords (30 nouns and 6 verbs). In this study, only the following 8 keywords, which were a subset of the word list and all nouns, were analyzed: 'sea-sheep', 'seat-sheet', 'cell-shell' and 'sack-shack'. This selection was done to enable comparability with the data reported for adult and child talkers in Romeo et al. (2013) for the casual condition.

2.3 Speech recordings

The picture elicitation task was run in a sound-treated booth using DMDX software [Forster and Forster, 2003], with participants wearing Beyerdynamic DT297PV microphone headsets, and the speech recorded at a sampling rate of 22,050 Hz. A picture appeared on the screen and, for nouns, participants were instructed to name each picture using the following frame sentence: 'I can see a (noun)'. The 36 pictures were each presented 8 times in a pseudo-randomized order, with nouns and verbs in separate blocks. In this first session, talkers were given the following instruction: 'when you say the sentences, try to speak casually as if talking to a friend'. In a following session, carried out on a separate day, the same items were recorded again but this time, talkers were told: 'when you say the sentences, try to speak very clearly as if you are talking to a person who is hearing-impaired'. Out of the 5120 tokens that were recorded for the subset of minimal pairs analyzed here, 145 tokens (or 2.8%) were removed from the analysis due to the sound file recorded via DMDX being truncated leaving 2477 and 2498 tokens for /s/ and /ʃ/ respectively, for analysis.

2.4 Speech analysis

The analysis method adopted in Romeo et al. (2013) is summarized here: see that paper for further details. Markers were placed in Praat [Boersma & Weenink, 2012] at the start of the frication portion and at the end of the frication portion excluding portions of mixed excitation. For each token, R scripts [Reidy, 2013] were used to band-pass filter the audio file then compute multitaper spectra using eight tapers for the middle 50% portion of the fricative; the four spectral moments were then obtained. Phoneme means and variance (calculated as the standard deviations of all tokens per talker) were obtained for three spectral features known to distinguish /s/ and /ʃ/ [Jongman et al., 2000]: fricative centroid, skewness and kurtosis. The centroid or center of gravity represents the mean frequency weighted by amplitude; fricative skewness represents the balance of energy between low and high frequency regions and kurtosis represents the peakedness of the energy distribution. To quantify phoneme discriminability, three additional measures were derived for each of these features per talker, as in Romeo et al. (2013). Between-category distance was calculated as the difference between the mean values for both tokens. Within-category dispersion was calculated as the mean of the standard deviations across /s/ and /ʃ/ tokens for each measure. Finally, the increased discriminability which could result from either strategy (increasing cross-category distance and decreasing within-category dispersion) was calculated as the difference between the mean values of two distributions (distance) times the square root of 2, divided by the square root of the sum of the within-category variances [d(a) measure; as discussed in Newman et al., 2001; Romeo et al., 2013]. The d(a) measure makes the assumption that the data are normally distributed: the between-category distance and within-category dispersion measures for centroid and skewness were normally distributed for both speaking conditions (Shapiro-Wilks, $p > .05$) but failed the test of normality as a results of positive skew for kurtosis. The data for kurtosis were then log transformed (to base 10) which normalized the data for both measures and conditions. The category discriminability index was re-calculated for kurtosis using the log-transformed data. Subsequently, all analyses were run for both untransformed and transformed data. Because parametric statistics are relatively robust to violations of normality, and transforming the data did not change the level of statistical significance, only analyses from the untransformed data is reported here.

3. Results

Means and standard deviations for fricative centroid, skewness and kurtosis for /s/ and /ʃ/ were first obtained for each talker and then calculated for male and female talker groups (see Table 1). These were then used to calculate cross-category distance, within-category dispersion and discriminability (see Table 2) as described above. Histograms for fricative centroids, skewness and kurtosis for /s/ and /ʃ/ are presented in Supplemental Materials.

A repeated-measures ANOVA was carried out with a within-subject factor of speaking style (casual, clear) and between-subject factor of talker sex on measures of cross-category distance and within-category dispersion for centroid, skewness and kurtosis.

Cross-category distance increased significantly in the clear relative to the casual condition for fricative centroid [$F(1,38)=16.27$; $p<.001$] but the effect of speaking style on cross-category distance was not significant for skewness [$F(1,38)=0.381$; $p=.541$] or kurtosis [$F(1,38)=1.63$; $p=.209$]. Distribution means for centroids were investigated further in a repeated-measures ANOVA with within-subject factors of consonant and speaking style, and between-subject effect of talker sex to establish whether the clear speech adaptations affected one consonant distribution more than the other. The effect of speaking style was significant [$F(1,38)=4.15$; $p=.049$] and there was a significant consonant by condition interaction [$F(1,38)=16.27$; $p<.001$]: this was due to a greater change in distribution means across speaking styles for /s/ (casual M = 7110; clear M = 7293) than /ʃ/ (casual M = 4570; clear M = 4543). For cross-category distance, the between-subject effect of talker sex was significant for centroid [$F(1,38)=22.28$; $p<.001$], skewness [$F(1,38)=18.74$; $p<.001$] and kurtosis [$F(1,38)=21.02$; $p<.001$] with, for all measures, a greater cross-category distance for women. The talker sex by condition interactions were non-significant for all measures.

A focus of the investigation was whether talkers produced less dispersed categories in their clear speaking style. The effect of condition on within-category dispersion was found to be non-significant for fricative centroid [$F(1,38)=.704$; $p=.407$] and for kurtosis [$F(1,38)=2.44$; $p=.127$] while dispersion decreased in the clear relative to conversational condition for skewness [$F(1,38)=5.41$; $p=.025$]. The effect of talker sex on dispersion was only significant for kurtosis [$F(1,38)=7.94$; $p=.008$], with women showing greater dispersion than men. The sex by condition interactions were non-significant for all measures.

Finally, we investigated whether the fricative contrast was more discriminable in the clear relative to the casual condition, using the d(a) measure of category discriminability. As units of measurement for (a) are directly comparable across centroid, skewness and kurtosis, they were analyzed in a single repeated-measures ANOVA with within-subject factors of cue (centroid, kurtosis, skewness) and condition (casual, clear) and between-subject factor of talker sex. The effect of cue [$F(2,76)=325.95$; $p<.001$] of talker sex [$F(1,38)=11.36$; $p=.002$] and the cue by talker sex interaction [$F(2,76)=18.19$; $p<.001$] were significant. Fricatives were significantly more discriminable in terms of their centroid values (M = 8.430) than their skewness (M = -2.521) or kurtosis (M=1.565) which also differed significantly from each other. Fricatives by female talkers (M = 2.931) were more discriminable than those by male talkers (M = 2.052), and this sex difference was greater for centroid and skewness than kurtosis measures. The main effect of condition was not significant [$F(1,38)=.628$; $p=.433$], and the interaction between cue and condition just failed to reach significance [$F(2,76)=3.598$; $p=.051$, Greenhouse-Geisser corrected]: discriminability therefore did not significantly increase in the clear condition across all three cues (M=2.544) relative to the casual condition (M=2.438).

As suggested by Buz et al. (2016), rather than an overall reduction in within-category dispersion, a useful strategy when producing clear speech would be to avoid producing potentially ambiguous fricatives in the region where they are more likely to overlap with another fricative category. We calculated the difference between the 95th percentile value for /ʃ/ and the 5th percentile for /s/ for each talker, which are the relevant 'tails' of the distribution (see Fig. 1). For the centroid measure, this distance between distribution tails increased significantly across the casual and clear styles [$F(1,38)=6.75$; $p=.013$], from 1886 Hz (SD = 720) to 2150 Hz (SD = 589) for women and from 1038 Hz (SD = 514) to 1206 Hz (SD = 680) for men. The effect of talker sex was also significant [$F(1,38)=24.34$; $p<.001$], with a greater distance between tails for women (M = 2018) than men (M = 1122) and no significant condition by sex interaction. For skewness, neither the effect of condition [$F(1,38)=.749$; $p=.392$] nor the condition by talker sex interaction [$F(1,38)=2.41$; $p=.129$] were significant, although overall the distance between tails in terms of skewness was greater for women (M = -2.838) than men (M = -2.009) [$F(1,38)=15.75$; $p=.000$]. A similar pattern was obtained for kurtosis: neither the effect of condition [$F(1,38)=.576$; $p=.453$] nor the condition by talker sex interaction [$F(1,38)=1.40$; $p=.244$] were significant. Overall the distance between tails in terms of kurtosis was smaller for women (M = -.681) than men (M = -2.579) [$F(1,38)=7.41$; $p=.01$].

These results suggest that, at least in terms of centroid frequency, a key marker of fricative place of articulation, talkers reduced the number of tokens produced in the ambiguous region for the /s/-/ʃ/ distinction. However, as seen in Fig. 1, it was not the case that all talkers used the strategy of increasing distance between distribution tails in the ambiguous region, as data points along the diagonal or in the lower half of the scatterplot represent talkers who show no difference in the distance between /ʃ/ and /s/ distribution tails.

Overall, therefore, in a picture elicitation task that included multiple randomized iterations of minimal pairs differing in initial /s/-/ʃ/, an instruction to speak clearly did not lead to a significant increase in the discriminability of the initial fricatives. Although the distance between the mean centroids increased in the clear condition, there was no effect of condition for the measures of skewness and kurtosis, and there was little evidence that talkers produced less dispersed categories as the effect of within-category dispersion was only significant for skewness. However, there was some evidence of adaptation in the clear condition as, on average, talkers produced fewer tokens with centroid values that could be in the ambiguous region for the /s/-/ʃ/ distinction; note that there was evidence of individual variability in the use of this strategy. There was a clear finding of greater cross-category distance in women for centroid, skewness and kurtosis leading to more discriminable categories for women than men for all three measures. Moreover, while Romeo et al. (2013) only investigated fricative centroid and showed a talker sex effect for this measure in both older children and adults, this study confirms and extends these finding by showing that a talker sex effect is also present for the measures of skewness and kurtosis. However, as in Maniwa et al. (2009), there were no significant sex by condition interactions suggesting that, while women had more discriminable fricative contrasts overall, they did not make greater segmental adaptations than men in the clear speech condition. Thus, despite the fact that women are perceived as more intelligible than men, they did not differ from men in their segmental adaptations in the clear speech condition.

4. Discussion

This study investigated whether talkers enhance segmental contrasts when instructed to speak clearly. There is now ample evidence that talkers do aim to enhance phonetic contrasts at a segmental level when tasks focus attention on a contrast between potentially confusable phonetic categories, either directly or indirectly. Talkers may achieve this by increasing the distance between phoneme categories but there is also evidence of changes in phonetic category structure to reduce the production of ambiguous tokens [Buz et al., 2016].

It has been argued by Buz et al., within their adaptive speaker framework, that these changes occur because talkers adapt their speech based on the perceived communicative success of their production. These adaptations take place even in the absence of explicit clarification requests when there is potential for confusion; the presence of interlocutor feedback could thus be seen as a requisite for these segmental-level adaptations to occur. This study differed from Buz et al. in two crucial points: there was no indication of communicative success available for the talker (i.e., no feedback present), and also, although the picture elicitation task involved randomized items from minimal pairs differing in initial /s/-/ʃ/ and /p/-/b/, each item was presented singly with no foils present (however cf. Buz & Jaeger, 2016 for implicit across-trial contrasting present in these types of elicitation tasks). Despite this, there was still evidence of some adaptations made at the segmental level to reduce potential ambiguity and increase perceived intelligibility [Maniwa, et al., 2008] in the clear speech condition, at least for the centroid measure which carries greater perceptual weight [Harris, 1958]. This was achieved by producing more distinct distributions and fewer ambiguous tokens but there was little evidence of producing less dispersed categories, even though compact distributions have been shown to lead to faster reaction times in speech perception tasks [Newman et al., 2001].

We have argued in previous work [e.g., Hazan & Baker, 2011] that talkers do make adaptations to their speech that are tailored to their interlocutor needs in various adverse conditions; these claims are in keeping with Buz et al.'s adaptive speaker framework. We would argue that the current results suggest that, whereas the presence of an interlocutor and of more realistic communicative behaviour may promote greater adaptation at the segmental level, this presence is not necessary for some adaptations to occur.

Acknowledgments

This work was funded by UK Economics and Social Research Council (RES-

062-23-0681). We acknowledge Rachel Baker who collected and analyzed the LUCID corpus. The LUCID corpus is available at the OSCAAR archive based at Northwestern University (https://oscaar.ci.northwestern.edu/).

References and links

Boersma, P. & Weenink, D. (**2012**). "Praat: doing phonetics by computer" (Version 5.3.16) (Computer program). Retrieved July 10, 2012, from http://www.praat.org/

Bradlow, A., Torretta, G., & Pisoni, D. (**1996**). "Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics," Speech Communication 20, 255-272.

Buz, E., & Jaeger, T.F. (**2016**). "The (in)dependence of articulation and lexical planning during isolated word production," Language, Cognition and Neuroscience 31, 404-424.

Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (**2016**). "Dynamically adapted context-specific hyper-articulation: Feedback from interlocutors affects speakers' subsequent pronunciations," Journal of Memory and Language 89, 68-86.

Cooke, M., King, S., Garnier, M., & Aubanel, V. (**2014**). "The listening talker: A review of human and algorithmic context-induced modifications of speech," Computer Speech and Language 28, 543-571.

Forster, K. I., & Forster, J. C. (**2003**). "DMDX: A windows display program with millisecond accuracy," Behavior Research Methods, Instruments, & Computers 35, 116-124.

Granlund, S., Hazan, V., & Baker, R. (**2012**). "An acoustic-phonetic comparison of the clear speaking styles of late Finnish-English bilinguals," Journal of Phonetics 40, 509-520.

Harris K. S. (**1958**). "Cues for the discrimination of American English fricatives in spoken syllables," Language and Speech 1, 1–7.

Hazan, V. & Baker, R. (**2011**). "Acoustic-phonetic characteristics of speech produced with communicative intent to counter adverse listening conditions," Journal of the Acoustical Society of America 130, 2139-2152.

Hazan, V., & Markham, D. (2004). "Acoustic-phonetic correlates of talker intelligibility for adults and children," Journal of the Acoustical Society of America 116, 3108-3118.

Jongman, A., Wayland, R., & Wong, S. (**2000**). "Acoustic characteristics of English fricatives," Journal of the Acoustical Society of America 108, 1252-1263.

Maniwa, K., Jongman, A., & Wade, T. (**2008**). "Perception of clear English fricatives by normal-hearing and simulated hearing-impaired listeners," Journal of the Acoustical Society of America 123, 1114–1125.

Maniwa, K., Jongman, A., & Wade, T. (**2009**). "Acoustic characteristics of clearly spoken English fricatives." Journal of the Acoustical Society of America 125, 3962–3973.

Newman, R. S., Clouse, S. A., & Burnham, J. L. (**2001**). “The perceptual consequences of within-talker variability in fricative production,” Journal of the Acoustical Society of America 109, 1181-1196.

Reidy, P. F. (**2013**). “An introduction to random processes for the spectral analysis of speech data,” Ohio State University Working Papers in Linguistics, 67-116.

Romeo, R., Hazan, V. & Pettinato, M. (**2013**). “Developmental trends and perceptual effects of intra-talker variability in consonant production,” Journal of the Acoustical Society of America 134, 3781-3792.

Smiljanić, R. & Bradlow, A. (**2008**). “Stability of Temporal Contrasts across Speaking Styles in English and Croatian.” Journal of Phonetics, 36, 91-113.

Table 1: Mean centroid values (in Hz) and standard deviation values in italics for female (N=20) and male (N=20) talkers in conversational and clear speaking styles. Measures are based on 32 measurements per consonant per condition on average.

	Centroid				Skewness				Kurtosis			
	/s/		/ʃ/		/s/		/ʃ/		/s/		/ʃ/	
Group	Cas.	Clear	Cas.	Clear	Cas.	Clear	Cas.	Clear	Cas.	Clear	Cas.	Clear
Female (*N*=20)	7946 (367)	8158 (410)	5016 (649)	5018 (510)	-.94 (.52)	-.94 (.54)	.56 (.58)	.39 (.38)	4.95 (2.51)	4.63 (2.32)	.67 (1.31)	.41 (.66)
Male (*N*=20)	6275 (632)	6429 (552)	4125 (483)	4070 (505)	.19 (.73)	.10 (.44)	.78 (.45)	.77 (.43)	1.99 (1.4)	1.35 (1.37)	.42 (1.06)	.55 (1.26)

Table 2: For /s/ and /ʃ/, cross-category distance, within-category dispersion, discriminability and distance between tails for centroid (in Hz), skewness and kurtosis in casual and clear speaking styles. Measures are based on 32 tokens per consonant per condition on average. Standard deviations in parentheses.

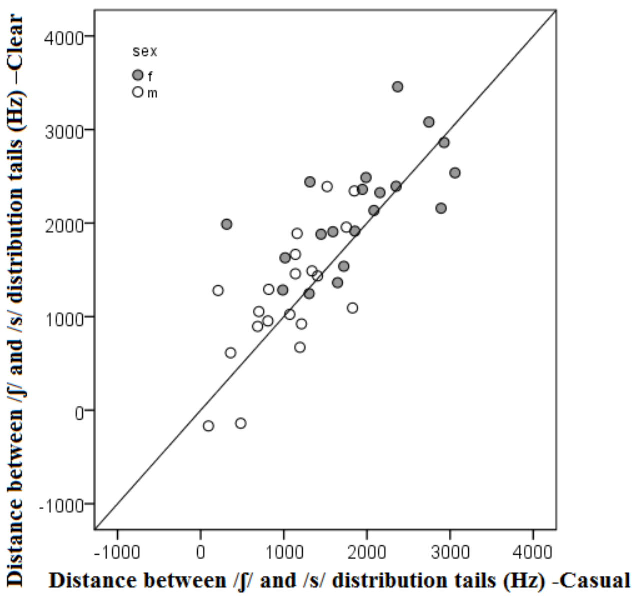
Talker	Centroid		Skewness		Kurtosis	
	Casual	Clear	Casual	Clear	Casual	Clear
	Cross-category distance					
Female	2930	3140	1.50	1.33	4.28	4.22
(N=20)	(671)	(567)	(0.67)	(0.51)	(3.00)	(2.68)
Male	2150	2360	0.59	0.67	1.57	0.80
(N=20)	(449)	(475)	(0.60)	(0.63)	(1.71)	(1.74)
	Within-category dispersion					
Female	313	306	0.44	0.39	1.95	1.64
(N=20)	(.60)	(67)	(0.12)	(0.09)	(1.09)	(0.59)
Male	332	365	0.42	0.39	1.29	1.21
(N=20)	(57)	(119)	(0.09)	(0.06)	(0.56)	(0.54)
	Discriminability					
Female	9.62	10.58	3.51	3.40	2.17	2.12
(N=20)	(3.18)	(3.31)	(1.80)	(1.60)	(1.34)	(0.98)
Male	6.52	6.99	1.49	1.69	1.31	0.66
(N=20)	(1.73)	(2.47)	(1.55)	(1.54)	(1.40)	(1.29)
	Distance between tails					
Female	1886	2150	-2.97	-2.71	-1.04	-0.20
(N=20)	(720)	(590)	(.92)	(.70)	(3.27)	(1.78)
Male	1038	1206	-1.97	-2.05	-2.49	-2.67
(N=20)	(514)	(680)	(.61)	(.70)	(2.87)	(2.47)

Figure captions

Fig. 1: Scatterplot showing the centroid value in Hz representing, for the casual (x axis) and clear speech (y axis) conditions, the difference between the 95th percentile value for the /ʃ/ distribution and the 5th percentile value for the /s/ distribution per talker.

Supplemental Materials: Histograms of fricative centroids (in Hz), Skewness and Kurtosis for productions of /s/ and /ʃ/ for Female (N=20) and Male (N=20) talkers and for Casual (top section of each graph) and Clear (bottom section of each graph) speaking conditions.

Fig 1.


sex
f
m
4000
3000
2000
1000
0
-1000
-1000
0
1000
2000
3000
4000
Distance between /ʃ/ and /s/ distribution tails (Hz) –Clear
Distance between /ʃ/ and /s/ distribution tails (Hz) -Casual