

Tuomainen , JASA-EL

1

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

4

5 Outi Tuomainen

6 Dept of Speech, Hearing and Phonetic Sciences, University College London
7 (UCL), Chandler House, 2 Wakefield Street, London WC1N 1PF, UK.

8 o.tuomainen@ucl.ac.uk

9

10 Valerie Hazan

11 Dept of Speech, Hearing and Phonetic Sciences, University College London
12 (UCL), Chandler House, 2 Wakefield Street, London WC1N 1PF, UK.

13 v.hazan@ucl.ac.uk

14

15 Rachel Romeo

16 Speech & Hearing Bioscience and Technology, Division of Medical Sciences,
17 Harvard University, TMEC 435, 260 Longwood Avenue, Boston MA 02115

18 rachelromeo@fas.harvard.edu

19

20

21

22

23 Running title: Phoneme category structure in clear speech

1 Abstract

2

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

14

15

16

17

18

19

20 © 2016 Acoustical Society of America

21

22 PACS number: 43.70.Mn

1 1. Introduction
2
3

4 When speakers are asked to speak clearly, either in an experimental
5 setting or when communicating with interlocutors who have a hearing
6 impairment or who are non-native, they make adaptations to various acoustic-
7 phonetic characteristics of their speech. Typically, speech produced in an
8 ‘instructed’ clear speaking style is slower, more intense, more hyperarticulated,
9 with greater pitch variations and contains more frequent pauses than speech
10 produced in a casual or conversational style [for a review, see Cooke et al.,
11 2014]. One issue of interest is the degree to which these clear speech adaptations
12 are aimed at making phonetic categories more easily discriminable rather than
13 just resulting in global enhancements.

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

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

6 A second strategy to increase distinctiveness would be to be more
7 consistent in phoneme production: This would reduce within-category
8 dispersion and potential category overlap. In perception experiments, slower
9 reaction times have been obtained for talkers who had greater within-category
10 variability, thus supporting the relevance of consistency in production for
11 speech perception [Newman et al., 2001]. In a recent study, when producing
12 target words containing a voiceless plosive (e.g., *pill*) for the benefit of an
13 interlocutor in the presence of potentially confusable foils containing a voiced
14 plosive (e.g., *bill*), talkers reduced the number of potentially confusable tokens
15 by reducing variance at the extremes of the category distribution, especially at
16 short VOTs [Buz et al., 2016]. In their ‘adaptive speaker framework’, Buz et
17 al. argued that the presence of interlocutor feedback was key in eliciting these
18 adaptations as the extent to which these clarifications occurred was dependent
19 on the communicative success of previous exchanges between the two talkers.
20 As within-category dispersion has received relatively little attention, the current
21 study investigated whether the production of a clear speaking style in the
22 absence of an interlocutor also involved such a strategy, or whether, on the

1 contrary, the effort to produce more distinct consonants led to increased
2 variability in production. The /s/-/ʃ/ fricative place contrast was chosen for
3 analysis. This phonetic contrast is primarily marked by spectral differences
4 between the two consonants, unlike stop voicing contrasts such as /p/-/b/ where
5 the primary cue, voice onset time, is a durational cue. With durational cues,
6 there are concerns about potential contrast enhancement effects being strongly
7 influenced by changes in articulation rates across conditions, as also noted in
8 Granlund et al. (2013). Fricatives produced in casual and clear styles have been
9 exhaustively analyzed in a study involving 500,000 measurements of fricative
10 tokens [Maniwa et al., 2009] but although that study considered a wide range
11 of acoustic cues to this contrast, it focused on cross-category distance and gave
12 less attention to within-category dispersion.

13 Here, we focus on the changes in internal category structure in adult
14 speech that result from adopting a clear speaking style. More specifically, the
15 aim of this study was to establish whether speech production in a clear speaking
16 style entails an increase in discriminability of phonetic contrasts, even in the
17 absence of genuine communicative behaviour that emerges when there is
18 feedback from the interlocutor. We also considered whether any increased
19 discriminability was as a result of greater internal consistency in the way in
20 which consonants are produced in a clear speaking style, thus producing less
21 dispersed categories as well as increasing category distance. As women have
22 been found in a number of studies to be more intelligible than men [Bradlow et

1 al.,1996; Hazan & Markham, 2004], we also examined the data for sex effects
2 to see whether women showed evidence of making greater segmental
3 adaptations in clear speech than men. The LUCID corpus is a large corpus of
4 spontaneous and read casual and clear speech [Hazan & Baker, 2011]. It
5 includes on average 32 iterations per consonant for each of 40 adult talkers for
6 the /s/-/ʃ/ fricative place contrast in word-initial position, obtained via picture
7 elicitation; thus, it enables us to examine within-category dispersion and across-
8 category distance in two ‘instructed’ speaking styles.

9

10 2. Method

11

12 2.1 *Talkers*

13

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

18

19 2.2 *Materials*

20

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

9

10 2.3 *Speech recordings*

11

12 The picture elicitation task was run in a sound-treated booth using DMDX
13 software [Forster and Forster, 2003], with participants wearing Beyerdynamic
14 DT297PV microphone headsets, and the speech recorded at a sampling rate of
15 22,050 Hz. A picture appeared on the screen and, for nouns, participants were
16 instructed to name each picture using the following frame sentence: ‘I can see
17 a (noun)’. The 36 pictures were each presented 8 times in a pseudo-randomized
18 order, with nouns and verbs in separate blocks. In this first session, talkers were
19 given the following instruction: ‘when you say the sentences, try to speak
20 casually as if talking to a friend’. In a following session, carried out on a
21 separate day, the same items were recorded again but this time, talkers were

1 told: ‘when you say the sentences, try to speak very clearly as if you are talking
2 to a person who is hearing-impaired’. Out of the 5120 tokens that were recorded
3 for the subset of minimal pairs analyzed here, 145 tokens (or 2.8%) were
4 removed from the analysis due to the sound file recorded via DMDX being
5 truncated leaving 2477 and 2498 tokens for /s/ and /ʃ/ respectively, for analysis.

6

7 2.4 *Speech analysis*

8

9 The analysis method adopted in Romeo et al. (2013) is summarized here: see
10 that paper for further details. Markers were placed in Praat [Boersma &
11 Weenink, 2012] at the start of the frication portion and at the end of the frication
12 portion excluding portions of mixed excitation. For each token, R scripts
13 [Reidy, 2013] were used to band-pass filter the audio file then compute
14 multitaper spectra using eight tapers for the middle 50% portion of the fricative;
15 the four spectral moments were then obtained. Phoneme means and variance
16 (calculated as the standard deviations of all tokens per talker) were obtained for
17 three spectral features known to distinguish /s/ and /ʃ/ [Jongman et al., 2000]:
18 fricative centroid, skewness and kurtosis. The centroid or center of gravity
19 represents the mean frequency weighted by amplitude; fricative skewness
20 represents the balance of energy between low and high frequency regions and
21 kurtosis represents the peakedness of the energy distribution. To quantify

1 phoneme discriminability, three additional measures were derived for each of
2 these features per talker, as in Romeo et al. (2013). Between-category distance
3 was calculated as the difference between the mean values for both tokens.
4 Within-category dispersion was calculated as the mean of the standard
5 deviations across /s/ and /ʃ/ tokens for each measure. Finally, the increased
6 discriminability which could result from either strategy (increasing cross-
7 category distance and decreasing within-category dispersion) was calculated as
8 the difference between the mean values of two distributions (distance) times the
9 square root of 2, divided by the square root of the sum of the within-category
10 variances [d(a) measure; as discussed in Newman et al., 2001; Romeo et al.,
11 2013]. The d(a) measure makes the assumption that the data are normally
12 distributed: the between-category distance and within-category dispersion
13 measures for centroid and skewness were normally distributed for both
14 speaking conditions (Shapiro-Wilks, $p > .05$) but failed the test of normality as a
15 results of positive skew for kurtosis. The data for kurtosis were then log
16 transformed (to base 10) which normalized the data for both measures and
17 conditions. The category discriminability index was re-calculated for kurtosis
18 using the log-transformed data. Subsequently, all analyses were run for both
19 untransformed and transformed data. Because parametric statistics are
20 relatively robust to violations of normality, and transforming the data did not
21 change the level of statistical significance, only analyses from the
22 untransformed data is reported here.

1

2 3. Results

3

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

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

14 Cross-category distance increased significantly in the clear relative to the casual
15 condition for fricative centroid [$F(1,38)=16.27$; $p<.001$] but the effect of
16 speaking style on cross-category distance was not significant for skewness
17 [$F(1,38)=0.381$; $p=.541$] or kurtosis [$F(1,38)=1.63$; $p=.209$]. Distribution
18 means for centroids were investigated further in a repeated-measures ANOVA
19 with within-subject factors of consonant and speaking style, and between-
20 subject effect of talker sex to establish whether the clear speech adaptations
21 affected one consonant distribution more than the other. The effect of speaking

1 style was significant [$F(1,38)=4.15$; $p=.049$] and there was a significant
2 consonant by condition interaction [$F(1,38)=16.27$; $p<.001$]: this was due to a
3 greater change in distribution means across speaking styles for /s/ (casual $M =$
4 7110; clear $M = 7293$) than /ʃ/ (casual $M = 4570$; clear $M = 4543$). For cross-
5 category distance, the between-subject effect of talker sex was significant for
6 centroid [$F(1,38)=22.28$; $p<.001$], skewness [$F(1,38)=18.74$; $p<.001$] and
7 kurtosis [$F(1,38)=21.02$; $p<.001$] with, for all measures, a greater cross-
8 category distance for women. The talker sex by condition interactions were non-
9 significant for all measures.

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

19 Finally, we investigated whether the fricative contrast was more discriminable
20 in the clear relative to the casual condition, using the $d(a)$ measure of category
21 discriminability. As units of measurement for (a) are directly comparable across

1 centroid, skewness and kurtosis, they were analyzed in a single repeated-
2 measures ANOVA with within-subject factors of cue (centroid, kurtosis,
3 skewness) and condition (casual, clear) and between-subject factor of talker sex.
4 The effect of cue [$F(2,76)=325.95$; $p<.001$] of talker sex [$F(1,38)=11.36$;
5 $p=.002$] and the cue by talker sex interaction [$F(2,76)=18.19$; $p<.001$] were
6 significant. Fricatives were significantly more discriminable in terms of their
7 centroid values ($M = 8.430$) than their skewness ($M = -2.521$) or kurtosis
8 ($M=1.565$) which also differed significantly from each other. Fricatives by
9 female talkers ($M = 2.931$) were more discriminable than those by male talkers
10 ($M = 2.052$), and this sex difference was greater for centroid and skewness than
11 kurtosis measures. The main effect of condition was not significant
12 [$F(1,38)=.628$; $p=.433$], and the interaction between cue and condition just
13 failed to reach significance [$F(2,76)=3.598$; $p=.051$, Greenhouse-Geisser
14 corrected]: discriminability therefore did not significantly increase in the clear
15 condition across all three cues ($M=2.544$) relative to the casual condition
16 ($M=2.438$).

17 As suggested by Buz et al. (2016), rather than an overall reduction in within-
18 category dispersion, a useful strategy when producing clear speech would be to
19 avoid producing potentially ambiguous fricatives in the region where they are
20 more likely to overlap with another fricative category. We calculated the
21 difference between the 95th percentile value for /f/ and the 5th percentile for /s/
22 for each talker, which are the relevant ‘tails’ of the distribution (see Fig. 1). For

1 the centroid measure, this distance between distribution tails increased
2 significantly across the casual and clear styles [$F(1,38)=6.75$; $p=.013$], from
3 1886 Hz ($SD = 720$) to 2150 Hz ($SD = 589$) for women and from 1038 Hz (SD
4 = 514) to 1206 Hz ($SD = 680$) for men. The effect of talker sex was also
5 significant [$F(1,38)=24.34$; $p<.001$], with a greater distance between tails for
6 women ($M = 2018$) than men ($M = 1122$) and no significant condition by sex
7 interaction. For skewness, neither the effect of condition [$F(1,38)=.749$;
8 $p=.392$] nor the condition by talker sex interaction [$F(1,38)=2.41$; $p=.129$] were
9 significant, although overall the distance between tails in terms of skewness was
10 greater for women ($M = -2.838$) than men ($M = -2.009$) [$F(1,38)=15.75$;
11 $p=.000$]. A similar pattern was obtained for kurtosis: neither the effect of
12 condition [$F(1,38)=.576$; $p=.453$] nor the condition by talker sex interaction
13 [$F(1,38)=1.40$; $p=.244$] were significant. Overall the distance between tails in
14 terms of kurtosis was smaller for women ($M = -.681$) than men ($M = -2.579$)
15 [$F(1,38)=7.41$; $p=.01$].

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

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

1 from men in their segmental adaptations in the clear speech condition.

2

3 4. Discussion

4

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

12 It has been argued by Buz et al., within their adaptive speaker framework, that
13 these changes occur because talkers adapt their speech based on the perceived
14 communicative success of their production. These adaptations take place even
15 in the absence of explicit clarification requests when there is potential for
16 confusion; the presence of interlocutor feedback could thus be seen as a
17 requisite for these segmental-level adaptations to occur. This study differed
18 from Buz et al. in two crucial points: there was no indication of communicative
19 success available for the talker (i.e., no feedback present), and also, although
20 the picture elicitation task involved randomized items from minimal pairs
21 differing in initial /s/-/ʃ/ and /p/-/b/, each item was presented singly with no foils

1 present (however cf. Buz & Jaeger, 2016 for implicit across-trial contrasting
2 present in these types of elicitation tasks). Despite this, there was still evidence
3 of some adaptations made at the segmental level to reduce potential ambiguity
4 and increase perceived intelligibility [Maniwa, et al., 2008] in the clear speech
5 condition, at least for the centroid measure which carries greater perceptual
6 weight [Harris, 1958]. This was achieved by producing more distinct
7 distributions and fewer ambiguous tokens but there was little evidence of
8 producing less dispersed categories, even though compact distributions have
9 been shown to lead to faster reaction times in speech perception tasks [Newman
10 et al., 2001].

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

18

19

Acknowledgments

20

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

1 062-23-0681). We acknowledge Rachel Baker who collected and analyzed the
 2 LUCID corpus. The LUCID corpus is available at the OSCAAR archive based
 3 at Northwestern University (<https://oscaar.ci.northwestern.edu/>).

4 **References and links**

- 5
- 6 Boersma, P. & Weenink, D. (2012). "Praat: doing phonetics by computer"
 7 (Version 5.3.16) (Computer program). Retrieved July 10, 2012, from
 8 <http://www.praat.org/>
- 9 Bradlow, A., Torretta, G., & Pisoni, D. (1996). "Intelligibility of normal speech
 10 I: Global and fine-grained acoustic-phonetic talker characteristics," *Speech*
 11 *Communication* 20, 255-272.
- 12 Buz, E., & Jaeger, T.F. (2016). "The (in)dependence of articulation and lexical
 13 planning during isolated word production," *Language, Cognition and*
 14 *Neuroscience* 31, 404-424.
- 15 Buz, E., Tanenhaus, M. K., & Jaeger, T. F. (2016). "Dynamically adapted
 16 context-specific hyper-articulation: Feedback from interlocutors affects
 17 speakers' subsequent pronunciations," *Journal of Memory and Language* 89,
 18 68-86.
- 19 Cooke, M., King, S., Garnier, M., & Aubanel, V. (2014). "The listening talker:
 20 A review of human and algorithmic context-induced modifications of speech,"
 21 *Computer Speech and Language* 28, 543-571.
- 22 Forster, K. I., & Forster, J. C. (2003). "DMDX: A windows display program

- 1 with millisecond accuracy,” Behavior Research Methods, Instruments, &
2 Computers 35, 116-124.
- 3 Granlund, S., Hazan, V., & Baker, R. (2012). “An acoustic-phonetic
4 comparison of the clear speaking styles of late Finnish-English bilinguals,”
5 Journal of Phonetics 40, 509-520.
- 6 Harris K. S. (1958). “Cues for the discrimination of American English fricatives
7 in spoken syllables,” Language and Speech 1, 1–7.
- 8 Hazan, V. & Baker, R. (2011). “Acoustic-phonetic characteristics of speech
9 produced with communicative intent to counter adverse listening conditions,”
10 Journal of the Acoustical Society of America 130, 2139-2152.
- 11 Hazan, V., & Markham, D. (2004). “Acoustic-phonetic correlates of talker
12 intelligibility for adults and children,” Journal of the Acoustical Society of
13 America 116, 3108-3118.
- 14 Jongman, A., Wayland, R., & Wong, S. (2000). “Acoustic characteristics of
15 English fricatives,” Journal of the Acoustical Society of America 108, 1252-
16 1263.
- 17 Maniwa, K., Jongman, A., & Wade, T. (2008). “Perception of clear English
18 fricatives by normal-hearing and simulated hearing-impaired listeners,”
19 Journal of the Acoustical Society of America 123, 1114–1125.
- 20 Maniwa, K., Jongman, A., & Wade, T. (2009). “Acoustic characteristics of
21 clearly spoken English fricatives.” Journal of the Acoustical Society of
22 America 125, 3962–3973.

- 1 Newman, R. S., Clouse, S. A., & Burnham, J. L. (2001). "The perceptual
2 consequences of within-talker variability in fricative production," Journal of
3 the Acoustical Society of America 109, 1181-1196.
- 4 Reidy, P. F. (2013). "An introduction to random processes for the spectral
5 analysis of speech data," Ohio State University Working Papers in Linguistics,
6 67-116.
- 7 Romeo, R., Hazan, V. & Pettinato, M. (2013). "Developmental trends and
8 perceptual effects of intra-talker variability in consonant production," Journal
9 of the Acoustical Society of America 134, 3781-3792.
- 10 Smiljanić, R. & Bradlow, A. (2008). "Stability of Temporal Contrasts across
11 Speaking Styles in English and Croatian." Journal of Phonetics, 36, 91-113.
- 12

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

5

6

	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)

7

8

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

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

5

1

2

Figure captions

3

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

8

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

13

14

15

16

17

18

19

20

1

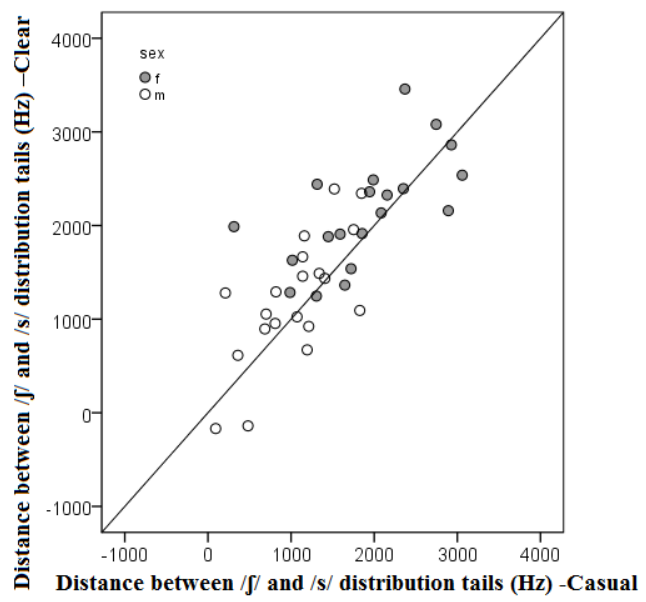
2

Fig 1.

3

4

5



6