

Perception of clear fricatives by normal-hearing and simulated hearing-impaired listeners

Kazumi Maniwa^{a)} and Allard Jongman

Department of Linguistics, The University of Kansas, Lawrence, Kansas 66044

Travis Wade

Posit Science Corporation, 225 Bush St. 7th floor, San Francisco, California, 94104

(Received 17 April 2007; accepted 12 November 2007)

Speakers may adapt the phonetic details of their productions when they anticipate perceptual difficulty or comprehension failure on the part of a listener. Previous research suggests that a speaking style known as clear speech is more intelligible overall than casual, conversational speech for a variety of listener populations. However, it is unknown whether clear speech improves the intelligibility of fricative consonants specifically, or how its effects on fricative perception might differ depending on listener population. The primary goal of this study was to determine whether clear speech enhances fricative intelligibility for normal-hearing listeners and listeners with simulated impairment. Two experiments measured babble signal-to-noise ratio thresholds for fricative minimal pair distinctions for 14 normal-hearing listeners and 14 listeners with simulated sloping, recruiting impairment. Results indicated that clear speech helped both groups overall. However, for impaired listeners, reliable clear speech intelligibility advantages were not found for non-sibilant pairs. Correlation analyses comparing acoustic and perceptual data indicated that a shift of energy concentration toward higher frequency regions and greater source strength contributed to the clear speech effect for normal-hearing listeners. Correlations between acoustic and perceptual data were less consistent for listeners with simulated impairment, and suggested that lower-frequency information may play a role. © 2008 Acoustical Society of America.

[DOI: 10.1121/1.2821966]

PACS number(s): 43.71.Es, 43.71.Ky, 43.71.Gv, 43.70.Mn [PEI]

Pages: 1114–1125

I. INTRODUCTION

Fricative consonants, especially non-sibilants, present considerable identification difficulty for hearing-impaired listeners and for normal-hearing listeners under adverse conditions (Boothroyd, 1984; Dubno and Levitt, 1981; Dubno *et al.*, 1982; Miller and Nicely, 1955; Owens, 1978; Owens *et al.*, 1972; Sher and Owens, 1974; Singh and Black, 1966; Soli and Arabia, 1979; Wang and Bilger, 1973). This study was designed to measure whether, and how, speakers may be able to alleviate this difficulty by deliberately producing fricatives more clearly.

A. Clear speech intelligibility advantage

Many language users spontaneously adapt the phonetic details of their speech on-line in response to social and communicative demands, adopting an intelligibility-enhancing speaking style when they anticipate or sense perceptual difficulty or comprehension failure on the part of a listener (due to, e.g., background noise, reverberation, hearing impairment, lack of linguistic/world knowledge). “Clear speech” has been elicited in laboratory settings (e.g., Bradlow and Bent, 2002; Bradlow *et al.*, 2003; Ferguson and Kewley-Port, 2002; Gagné *et al.*, 1994, 1995, 2002; Helfer, 1997, 1998; Iverson and Bradlow, 2002; Krause and Braid, 2002;

Liu *et al.*, 2004; Payton *et al.*, 1994; Picheny *et al.*, 1985; Schum, 1996; Uchanski *et al.*, 1996), and shown to result in intelligibility advantages relative to “conversational” speech ranging from 7 to 38 percentage points. Clearly spoken sentences benefit young normal-hearing listeners in noise and/or reverberation (Bradlow and Bent, 2002; Gagné *et al.*, 1995; Krause and Braid, 2002; Payton *et al.*, 1994; Uchanski *et al.*, 1996) and with simulated hearing loss or cochlear implants (Iverson and Bradlow, 2002; Liu *et al.*, 2004), hearing-impaired listeners in quiet (Picheny *et al.*, 1985; Uchanski *et al.*, 1996) and in noise or reverberation (Payton *et al.*, 1994; Schum, 1996), cochlear-implant users (Iverson and Bradlow, 2002; Liu *et al.*, 2004), elderly listeners with or without hearing loss (Helfer, 1998; Schum, 1996), children with or without learning disabilities (Bradlow *et al.*, 2003) and (perhaps to a lesser extent) non-native listeners (Bradlow and Bent, 2002).

Recent results from Ferguson and Kewley-Port (2002) question the robustness of the “clear speech effect” and suggest that hyperarticulation strategies may interact in complicated ways with different types of signal degradation. While Ferguson and Kewley-Port found intelligibility benefits for clearly produced vowels with young, normal-hearing listeners, they observed *negative* clear-speech intelligibility benefits (better recognition of conversational tokens) with elderly hearing-impaired listeners, at least for one talker’s productions. This pattern was mostly due to front vowels. A hallmark of clear speech is a greater concentration of energy

^{a)}Author to whom correspondence should be addressed. Electronic mail: kazumi.maniwa@uni-konstanz.de

in higher frequencies, in terms of both overall spectral distributions and individual formant frequencies (e.g., Krause and Braida, 2002; Picheny *et al.*, 1985). Since F2 values for front vowels fell in a frequency region where these listeners had sloping hearing loss (above 2000 Hz), clear vowels' higher F2 resonances, on average, fell in regions of greater impairment than those of conversational vowels.

It is of course unclear whether the patterns observed for this talker are unique to him or whether they are typical of the production, and perception by hearing-impaired or older listeners, of clear front vowels. The present study was designed to determine whether clear speech advantages occur for another class of sounds with a preponderance of high-frequency energy, fricatives, over a range of talkers and for young normal-hearing listeners and listeners with simulated high-frequency hearing loss.

B. Talker-related acoustic correlates of clear speech intelligibility

A secondary goal of this study was to determine which aspects of clear fricative production influence intelligibility. Previous investigations of the intelligibility of clear and conversational speech that have included more than a single talker have revealed considerable differences in the magnitude of the clear speech effect across talkers (e.g., Bradlow *et al.*, 2003; Chen, 1980; Ferguson, 2002, 2004; Gagné *et al.*, 1994, 1995; Schum, 1996). A few studies have attempted to identify talker-specific acoustic-phonetic parameters that may be responsible for the clear speech effect, by relating intelligibility differences to acoustic differences in clear and conversational speech. The talker in Bradlow *et al.* (2003) who showed the greater intelligibility advantage for clear speech substantially decreased her speaking rate with increased frequency and duration of pauses. Ferguson (2002) compared ten vowel measurements (five steady-state metrics, four dynamic metrics, and duration) across speakers and found that "big benefit" talkers showed the greatest increases in front vowel F2, F1 range, and the overall size of the vowel space.

The present study included an extended analysis of this type. Intelligibility was tested using a database of 8800 clear and conversational fricative productions by 20 talkers (10 M, 10 F), for which several spectral, temporal, and amplitude measurements have been reported (Maniwa *et al.*, submitted). These fricatives were all produced in vowel-consonant-vowel (VCV) (/a/-fricative-/a/) contexts. Based on features known to contribute to the perception of fricatives (described in the next section), the following measurements were made for these sounds: the frequency of the peak in the discrete Fourier transform (DFT); the mean, standard deviation, skewness, and kurtosis of the (DFT) spectral distribution; F2 onset transitions; the slope of the power spectrum below and above the (expected) peak location; the mean fundamental frequency (f0) of the adjacent vowels; root-mean-square (rms) fricative amplitude relative to the surrounding vowels; *frequency-specific relative amplitude* (FSRA), i.e., the amplitude of the frication relative to the surrounding vowels in the F3 region for sibilants and the F5 region for non-sibilants; fricative harmonic-to-noise ratio (HNR), energy below

500 Hz, and duration. Very briefly, this analysis revealed several overall acoustic-phonetic modifications in the production of clear fricatives. Some of these effects were straightforwardly predictable based on previous findings (e.g. longer duration; energy at higher frequencies shown by higher peak, mean, and F2 frequencies; lower skewness indicating more positive spectral tilt; and steeper spectral slopes suggesting more defined peaks), and some were more surprising (especially lower relative amplitude). In most cases there were also Style \times Fricative interactions indicating that these effects differed depending on the fricative; these effects were usually in a direction such that the acoustic distance between neighboring sounds was maximized in clear speech. For all measures, there was a wide range of variability across talkers in the extent to which a modification was implemented. In the present study, correlation analyses of acoustic and intelligibility measures across talkers were performed to assess the contributions of different acoustic modifications to intelligibility.

C. Cues to English fricative identity in different listener groups

Acoustic cues that have been reported to affect perception of English fricative place of articulation for listeners with normal hearing include frication duration, spectrum, and amplitude, as well as adjacent formant transitions and vowel quality. Experiments using natural (Harris, 1958; Zeng and Turner, 1990), synthetic (Heinz and Stevens, 1961; Zeng and Turner, 1990), and hybrid (Nittrouer, 1992, 2002; Nittrouer and Miller, 1997a and 1997b) speech suggest that spectral cues are important for distinguishing sibilants, while formant transition cues may help to distinguish nonsibilants (Harris, 1958; Heinz and Stevens, 1961; Nittrouer, 2002) and take on more weight when spectral cues are ambiguous (Hedrick, 1997; Hedrick and Carney, 1997; Hedrick and Ohde, 1993; Hedrick and Younger, 2003; Whalen, 1981). Overall noise duration and amplitude seem to have less perceptual significance (Behrens and Blumstein, 1988; Hedrick, 1997; Hedrick and Carney, 1997; Hedrick and Ohde, 1993; Hughes and Halle, 1956; Jongman, 1989; cf. Guerlekian, 1981; McCasland, 1979a and 1979b), but manipulation of frication amplitude in particular frequency regions does influence listeners' perception of place of articulation for /s/-/ʃ/ and /s/-/θ/ contrasts (Hedrick, 1997; Hedrick and Carney, 1997; Hedrick and Ohde, 1993; Hedrick and Younger, 2003; Stevens, 1985). Fewer studies have investigated which acoustic components serve to distinguish voiced and voiceless fricatives. It appears that noise duration, the amplitude and duration of glottal vibration at the edge of the fricative, and the extent of F1 transitions interact in determining listener judgments of voicing for intervocalic fricatives (Stevens *et al.*, 1992).

Listeners do not seem to process fricative acoustic cues independently, but integrate information obtained from multiple dimensions; furthermore, the perceptual weights assigned to different acoustic properties depend on contexts and listeners. Adult listeners with normal hearing seem to make more use of spectral cues for place of articulation information (Heinz and Stevens, 1961; Harris, 1958; Hedrick and Ohde, 1993; Hughes and Halle, 1956; Nittrouer, 1992;

Nittrouer and Miller, 1997a and 1997b; Nittrouer, 2002; Zeng and Turner, 1990), and temporal information for the voicing distinction (Cole and Cooper, 1975; Raphael, 1972; Soli, 1982). Hearing-impaired listeners may have difficulty integrating amplitude and spectral cues, and may generally place less weight on formant transitions than listeners with normal hearing (Hedrick, 1997; Hedrick and Younger, 2003; Zeng and Turner, 1990). In addition, listeners with sloping hearing loss commonly have elevated thresholds, and reduced dynamic range, in regions relevant to fricative perception (e.g., Dubno *et al.*, 1982; Owens *et al.*, 1972; Sher and Owens, 1974). It is likely, then, that clear speech alternations involving fricative spectra may have different results depending on the listener population. To address this possibility, this study examined the perception of clear and conversational fricatives by normal-hearing listeners (Experiment 1) and listeners with simulated hearing impairment (Experiment 2).

D. Hypotheses

Two experiments were performed to address three questions. First, are clearly produced fricatives more intelligible than conversational fricatives for listeners with normal hearing in degraded conditions? Based on previous findings, we hypothesized that they would be, although the effects might vary depending on fricatives (e.g. Ferguson and Kewley-Port, 2002). Second, what acoustic modifications are related to intelligibility? It was hypothesized that not all strategies employed by talkers serve to improve fricative identification, although it was difficult to predict which modifications would be most effective given previous conflicting results. Third, do clear-speech intelligibility differences differ based on listener population, in particular for listeners with sloping hearing losses? We expected that hearing loss might interact with clear-speech strategies, perhaps resulting in reduced benefit where high-frequency information was critical.

II. EXPERIMENT 1: EFFECT OF CLEAR SPEECH FOR FRICATIVE RECOGNITION BY LISTENERS WITH NORMAL HEARING

A. Method

1. Participants

Fourteen normal-hearing listeners (8 F, 6 M) aged 19–32 were recruited from the University of California, Berkeley. Participants were native speakers of American English, without noticeable regional dialects. Participants reported normal hearing and no history of speech or language disorders. Listeners were paid for their participation in the experiment.

2. Materials

As discussed in Sec. I B, intelligibility was assessed using a previously described corpus of VCV stimuli (Maniwa *et al.*, submitted). Briefly, conversational and clear tokens were elicited using an interactive program that ostensibly attempted to identify the sequence of fricatives produced by a speaker. The program made frequent, systematic errors involving voicing and place alternations, after which the speaker repeated a sound more clearly, as if trying to disambiguate the production for an elderly or hearing-impaired listener. All stimuli were normalized to the same long-term (word-level) rms amplitude and presented at 60 dB sound pressure level using MATLAB (The Math Works, Inc., 2000). Test stimuli were delivered in a background of 12-talker (6 F, 6 M) babble recorded at a sampling rate of 44.1 kHz. A total of 60 s of babble was created for the purposes of the experiment; for each stimulus, a segment of babble was selected from a random location within this 60 s sample. The duration of this segment exceeded that of the test item by a total of 600 ms, with the test stimulus centered temporally in the babble. There were 5 and 100 ms linear on-off ramps for the target stimulus and the noise, respectively.

biguate the production for an elderly or hearing-impaired listener. All stimuli were normalized to the same long-term (word-level) rms amplitude and presented at 60 dB sound pressure level using MATLAB (The Math Works, Inc., 2000). Test stimuli were delivered in a background of 12-talker (6 F, 6 M) babble recorded at a sampling rate of 44.1 kHz. A total of 60 s of babble was created for the purposes of the experiment; for each stimulus, a segment of babble was selected from a random location within this 60 s sample. The duration of this segment exceeded that of the test item by a total of 600 ms, with the test stimulus centered temporally in the babble. There were 5 and 100 ms linear on-off ramps for the target stimulus and the noise, respectively.

3. Procedures and apparatus

The perception test employed a two-alternative forced-choice identification task. The eight fricatives were divided into eight minimal pairs, depending on place of articulation and voicing: /f/-/θ/, /v/-/ð/, /s/-/ʃ/, /z/-/ʒ/, /t/-/v/, /θ/-/ð/, /s/-/z/, and /ʃ/-/ʒ/. Each pair was tested separately for clear and conversational styles, for a total of 16 sub tests. Sub-test order was randomized across subjects in a single 1 h session. Subjects listened to stimuli presented via Koss headphones in sound-attenuated rooms, seated in front of a computer monitor and mouse. On each trial, test VCV and babble waveforms were scaled based on the selected signal-to-noise ratio (snr) (described below) and the constant target stimulus level, combined additively, and presented diotically to the subjects, who were prompted to identify the fricative from a minimal pair by using the mouse to click one of two letter combinations on the computer screen. Response alternatives were written: “ff,” “th,” “ss,” “sh,” “vv,” “dh,” “zz,” and “zh.” Listeners were first oriented to the spelling of response alternatives and the test procedure, and a ten-trial block of fricative tokens at a high snr (+10 dB) was run with feedback before each sub-test.

The goal of each sub test was to determine the snr threshold at which a distinction could be made with 75% accuracy. In each test, two 40-trial adaptive tracks were initiated at +3 dB and -3 dB snr and interleaved at random over the 80-trial block. Signal-to-noise ratio values for each track were selected using a Bayesian adaptive algorithm (ZEST; e.g., King-Smith *et al.*, 1994). The final threshold estimate was simply taken as the average (in dB) of the snr values for each track on the final (40th) trial. While this approach may have resulted in less precise measurements of thresholds that were further from the initial guesses (since termination was not based on confidence criteria) it was considered more important that participants were exposed to equal numbers of stimuli from each contrast pair; ± 20 dB were chosen as absolute maximum and minimum allowable snr values. Individual test tokens were selected randomly from the productions of the 20 speakers, so that speakers and productions would, on average, be represented with equal frequency.

4. Data analysis

The clear speech intelligibility effect was tested using a repeated measures analysis of variance (ANOVA) with two within-subject factors (Style; two levels, Pair; eight levels) and threshold (dB snr) as the dependent variable. In order to assess the effect of pair type more thoroughly, another repeated measures ANOVA was calculated with three within-subject factors. One of the factors was Style. The second factor, Sibilance, depended on whether the pair consisted of sibilant fricatives or non-sibilant fricatives, and the third factor, Distinction, was labeled depending on whether the pair involved a place or voicing distinction. Pairwise comparisons for significant within-subject factors were done using Bonferroni corrected 95% confidence intervals.

In addition, as a first step in determining which acoustic modifications were related to intelligibility, correlation analyses were carried out across the 20 speakers included in the experiment, relating differences in their production strategies to differences in their clear-speech benefit. First, for each speaker, a single clear-minus-conversational difference value, averaged over all fricatives and productions, was calculated for each of the 14 acoustic measures reported in the Maniwa *et al.* (submitted) study: DFT peak location (1), the first four spectral moments moments (M1–M4; 2–5), F2 onset transitions (6), spectral slopes below (7) and above (8) typical peak locations (SlpBef, SlpAft), averaged f0 of adjacent vowels (9), normalized rms amplitude (rmsamp, 10), frequency-specific relative amplitude (FSRA) (11), HNR (12), energy below 500 Hz (13) and fricative duration (14). For (1)–(5), (7)–(8), (10), and (13), analyses considered 40 ms Hamming windowed segments at five locations: centered over the fricative onset, 25, 50, and 75% points, and offset (window (W) 1–5). For (6), acoustic values were derived at fricative onset and offset and each vowel midpoint from an analysis (W1–4). For (9), f0 was averaged across the vowels preceding and following the target. For (11), (12) and (14), the values were obtained over the entire fricative. In the present analyses, 50-order linear predictive coding (LPC) peaks (at the same five locations) were included as well, and f0 was considered separately preceding and following the fricative. Thus, the total number of acoustic values considered was 59. Since many of these variables were closely related and correlated strongly, principal component analysis (PCA) was used to transform the data (equated for mean and variance) into a smaller number of more independent dimensions, which were also compared with talkers' clear speech intelligibility benefits.

Next, a similar overall clear-minus-conversational *intelligibility* difference had to be estimated for each speaker. This was less straightforward, since the adaptive procedure ensured that overall accuracy (at least toward the end of sub tests) was about the same across fricative pairs and speaking styles. However, since different trials within sub tests involved different speakers and productions, absolute difficulty was not necessarily exactly equal for all stimuli with a given snr. This lack of homogeneity (which is inevitable when using natural productions) probably added some noise to the threshold estimation procedure. However, we were able to exploit it in order to measure, in parallel, differences in the

clear speech benefit across talkers. First, we verified that over the 32 total adaptive tracks that each listener in Experiment 1 heard, tokens from different speakers occurred, on average, with equal frequency and at equal signal-to-noise ratios. Then we simply took the clear-minus-conversational difference in accuracy (% correct), averaged across listeners, sub tests, and snr values, for each speaker as that speaker's approximate clear speech intelligibility advantage. While listener, sub test, and snr certainly all contributed to mean accuracy, we assumed that these contributions would essentially amount to random variability across speakers (serving only to make our measure of intelligibility advantage more conservative), and therefore no corrections were made based on these variables.

Of course, this comparison was limited in the types of acoustic-perceptual relationships it could detect. As reported by Maniwa *et al.* (submitted), clear fricatives were characterized not only by overall differences in acoustic measures depending on speaking style, but by numerous and complex Style \times Fricative interactions. Since correlation analysis capable of capturing these higher-order acoustic differences was not feasible given the constraints of the perception experiments described here (individual speakers were not represented well enough within subtests to ensure equalized average snr), we did not consider these patterns in the present study.

B. Results and discussion

1. Fricative intelligibility for listeners with normal hearing

Figure 1 shows mean snr thresholds as a function of fricative pair and speaking style. The Style \times Pair ANOVA showed an effect of Style [$F(1, 13)=149.5, p>0.001$], with 3.1 dB lower thresholds for clear speech than for conversational speech, indicating that clearly produced fricatives are more intelligible than casually produced fricatives for listeners with normal hearing in degraded listening conditions. The Pair effect was also significant [$F(7, 91)=113.8, p<0.001$]; across speaking styles, thresholds were lowest for the voiceless sibilant place of articulation contrast /s-/ʃ/, followed by /s-/z/ and /j-/ʒ/. Non-sibilant place of articulation pairs /f/-θ/ and /v/-ð/ were the most difficult, in accordance with previous studies (e.g. Jongman *et al.*, 2000; Miller and Nicely, 1955; Wang and Bilger, 1973). The Style \times Pair interaction was marginally significant [$F(7, 91)=2.1, p=0.051$], probably due to pairs /v/-ð/ and /f/-v/. *Post-hoc* comparisons revealed that the “clear speech effect” did not reach significance for these two pairs; all other pairs showed significant clear speech advantages.

The Style \times Sibilance \times Distinction Type ANOVA revealed a main effect of Sibilance [$F(1, 27)=370.9, p<.001$] with lower thresholds for sibilants than for non-sibilants. The main effect of Distinction Type was also significant [$F(1, 27)=103.7, p<0.001$] with lower thresholds for voicing distinctions relative to place of articulation distinctions. A Style \times Sibilance interaction [$F(1, 27)=10.33, p<0.01$] showed that while both sibilants and non-sibilants were more intelligible in clear speech, the effect was larger for sibilant

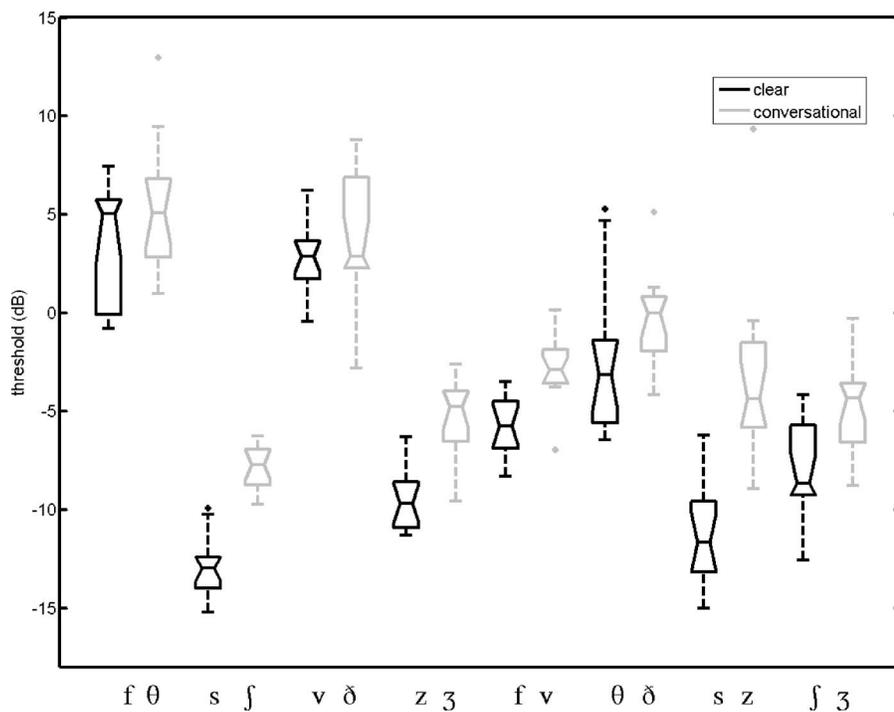


FIG. 1. Signal-to-noise ratio (snr) thresholds (dB) as a function of style and fricative pair in Experiment 1. Boxplots show the median, upper, and lower quartile, and outlier data (asterisks).

pairs. The Style \times Distinction Type interaction was not significant [$F < 1$]; clear speech resulted in similar benefits for place and voicing distinctions. There was no Style \times Sibilance \times Distinction Type interaction. In accordance with previous findings (e.g. Jongman *et al.*, 2000; Miller and Nicely, 1955; Wang and Bilger, 1973), sibilant pairs and voicing distinction pairs were easier to identify relative to non-sibilant and place of articulation pairs, respectively, regardless of speaking style.

2. Talker-related acoustic-phonetic correlates of clear intelligibility advantage

On average, individual talkers appeared in 336 (std. 18.8) clear and 336 (20.01) conversational trials. Due to the adaptive procedure and the initial threshold guesses of ± 3 dB across styles and pairs, talkers appeared at -4.91 dB (std. 0.28) and -2.61 dB (0.22) snr values, and were responded to with 81.9% (std. 4.58) and 77.3% (5.0) accuracy in clear and conversational conditions, respectively. Averaged across listeners, contrasts, and snr values, the clear-minus-conversational difference in accuracy (% correct) varied considerably across speakers, from -4% to $+11\%$ (mean 4.6%, std. 3.9%), at least partly as a result of differences in the clear speech strategies that these talkers employed (i.e., this difference did not correlate well [$p=0.34$] with clear-minus-conversational snr differences). As described above, then, individual speakers' previously reported average style-related differences in production were compared with their style-related intelligibility differences in an effort to relate clear speech benefits to specific acoustic modifications. Table I summarizes the results of Pearson's correlations between each individual acoustic measure and the clear-minus-conversational threshold difference. Positive correlations were obtained between intelligibility advantages and acoustic modifications in DFT and LPC peak location, spectral mo-

ment 1, and the slope before the peak, at most window locations. These results suggest that a shift of spectral energy to higher frequency regions and greater source strength (Jesus and Shadle, 2002) in clear fricatives—resulting in higher peak locations, higher frequency content on average, and more defined peaks—are most closely related to the overall intelligibility enhancement.

Principal components analysis of acoustic measures supported this observation. Figure 2 shows the contributions of individual acoustic measures to the first two components. The first component accounted for 41% of the variability; acoustic variables with the highest-magnitude coefficients for this component were those related to source strength and energy at higher frequencies (higher peaks and mean frequencies, lower skewness at central window locations). Talker scores for the first component correlated significantly with their clear speech benefit ($r=0.45$; $p=0.047$). The next two-components accounted for 14% and 11% of the variability, respectively. Intensity measures seemed to contribute most to the second component, and slope after peak locations most to the third; neither correlated significantly with the clear speech benefit ($p > 0.5$).

Since perception of place and voicing distinctions probably involve different acoustic cues, this analysis was repeated separately for the four place distinction subtests and the four voicing sub tests in the experiment. Comparison of these analyses suggested that most of the effects mentioned above were due to place of articulation distinctions. As shown in Table I, considering only place distinctions, strong positive correlations between clear speech intelligibility and acoustic differences were found for peak locations and slope before the peak, whereas negative correlations were seen for M3. Similarly, the first acoustic principal component correlated strongly with the clear speech benefit in place of articulation (POA) distinctions ($r=0.61$, $p=0.004$), but none of the

TABLE I. Correlation coefficients (Pearson's r) showing the relation between the clear-minus-conversational differences in acoustic measures and the clear-minus-conversational differences in the intelligibility (percent identification correctness) in Experiments 1 and 2. Significant values, $p < 0.001$, $p < 0.01$, and $p < 0.05$ are starred as ***, **, and *, respectively. Moderate values, $p < 0.1$ are marked as., and no effect was given N. Negative correlation was marked as ξ .

	Experiment 1			Experiment		
	Overall	Place	Voicing	Overall	Sibilant	Nonsibilant
Durs	0.3	0.11	0.25	-0.09	-0.02	-0.14
F2W1	-0.25	-0.21	-0.12	0.32	0.25	0.14
F2W2	0.34	0.45*	0.07	-0.28	-0.2	-0.19
F2W3	0.16	0.1	0.15	-0.33	-0.14	-0.43
F2W4	0.24	-0.08	0.34	0.05	-0.1	0.1
DFTpkW1	0.24	0.23	0.12	0.18	-0.11	0.25
DFTpkW2	0.53*	0.55*	0.21	0.1	-0.31	0.26
DFTpkW3	0.47*	0.56**	0.11	-0.03	-0.39	0.29
DFTpkW4	0.38	0.64**	-0.06	-0.03	-0.43	0.35
DFTpkW5	-0.12	-0.37	0.16	-0.08	-0.17	-0.06
HNR	-0.04	-0.19	0.08	-0.21	0.14	-0.3
Int500W1	0.01	0.27	0.22	-0.02	0.12	-0.14
Int500W2	-0.32	-0.37	-0.07	-0.01	-0.07	-0.08
Int500W3	-0.32	-0.42	-0.02	-0.16	-0.09	-0.3
Int500W4	-0.35	-0.39	-0.09	-0.16	-0.12	-0.25
Int500W5	0.26	0.32	0.06	0.43	0.41	0.44
M1W1	0.32	0.38	0.1	0.16	-0.13	0.24
M1W2	0.51*	0.41	0.29	0.17	0.07	0.27
M1W3	0.43	0.39	0.21	0.22	-0.05	0.39
M1W4	0.41	0.46*	0.12	0.14	-0.13	0.33
M1W5	0.17	-0.03	0.24	-0.01	-0.44	-0.02
M2W1	0.36	0.52*	0.04	0.02	-0.12	0.08
M2W2	0.25	0.26	0.1	0.06	-0.34	-0.16
M2W3	0.2	0.16	0.12	-0.06	-0.34	0.09
M2W4	0.14	0.2	0.02	-0.23	-0.51*	0.01
M2W5	0.19	-0.12	0.34	-0.19	-0.43	-0.25
M3W1	-0.36	-0.56**	-0.01	-0.15	-0.01	-0.16
M3W2	-0.46*	-0.56*	-0.11	-0.12	0.1	-0.14
M3W3	-0.36	-0.48*	-0.04	-0.09	0.34	-0.21
M3W4	-0.28	-0.54*	0.09	0.04	0.44	-0.14
M3W5	-0.38	-0.26	-0.26	-0.14	0.24	-0.08
M4W1	-0.31	-0.6**	-0.09	-0.11	-0.05	-0.06
M4W2	-0.31	-0.56*	-0.07	-0.01	0.24	0.02
M4W3	-0.24	-0.47*	-0.09	0.08	0.33	0.02
M4W4	-0.17	-0.49*	-0.19	0.21	0.55*	0.07
M4W5	-0.39	-0.28	-0.25	-0.11	0.26	-0.02
folIF0	0.24	0.66**	-0.23	0.39	0.15	0.49*
prevF0	-0.15	-0.1	-0.13	-0.11	-0.13	0.18
FSRA	0.04	-0.02	0.1	-0.09	-0.03	0.11
rmsampW1	0	-0.34	0.28	-0.18	0.1	-0.34
rmsampW2	-0.12	-0.33	0.15	-0.13	0.13	-0.27
rmsampW3	-0.06	-0.25	0.23	-0.21	0.14	-0.41
rmsampW4	-0.05	-0.24	0.17	-0.11	0.27	-0.32
rmsampW5	0.34	0.34	0.16	0.33	0.42	0.26
SlpAftW1	-0.42	-0.12	-0.4	-0.21	-0.03	-0.25
SlpAftW2	-0.29	-0.36	-0.04	-0.12	-0.09	-0.16
SlpAftW3	-0.16	-0.33	0.09	-0.01	0.22	-0.13
SlpAftW4	-0.02	-0.24	0.2	-0.1	0.27	-0.2
SlpAftW5	-0.15	-0.29	0.09	-0.27	-0.11	-0.42
SlpBefW1	0.42	0.54*	0.11	0.09	0.05	-0.06
SlpBefW2	0.39	0.54*	0.04	0	0.17	-0.12
SlpBefW3	0.39	0.62**	-0.01	0.01	0.28	-0.24
SlpBefW4	0.45*	0.62**	0.06	0.02	0.23	-0.15
SlpBefW5	0.29	0.45*	-0.01	0.12	0.34	-0.14
LPCPeakW1	0.3	0.35	0.11	0.15	0.01	0.17

quency ranges where many impaired listeners have elevated thresholds, and since Experiment 1 suggests that this may be increasingly so for clear fricatives. It is possible that listeners with sloping hearing loss cannot make use of enhanced acoustic-phonetic information since it is less audible to them. To assess how this aspect of hearing impairment would affect the perception of clear fricative sounds, we repeated the perception experiment using stimuli processed to simulate sloping hearing loss.

2. Implementation

Sloping, recruiting hearing loss was simulated in a manner similar to that described by Moore and Glasberg (1993), with some modifications due to a higher sampling rate (44.1 kHz) and the fact that all processing was done on-line during the experiment. Following the combination of signal and noise components, stimuli were separated into 24 equivalent rectangular bandwidth (ERB)-spaced bands, from 100 Hz to 22.05 kHz, using fourth order gammatone filters (Slaney, 1998). For each band, a smoothed envelope (E) was derived by low-pass filtering the full-wave rectified waveform at 100 Hz (fourth order Butterworth filter, implemented in both forward and reverse directions to minimize phase distortions). The temporal fine structure for the band was then extracted by dividing the original waveform by this envelope. Loss simulation was accomplished by raising the envelope to a power related to the slope of the loudness growth function:

$$E_p = E^N,$$

where N is frequency dependent. Following Moore and Glasberg (1993), N was a constant 1.5 at bands up to 900 Hz, increased linearly to 3.0 at 4500 Hz, and remained at this value for all higher bands. Finally, the modified stimulus was obtained by multiplying E_p by the fine structure and summing the resulting band-limited waveforms. All processing was performed in MATLAB. Processing on average took ~ 2 s; this resulted in an inter-trial interval that a few participants found slightly annoying but generally not distracting.

B. Experiment method

1. Participants

Fourteen normal-hearing listeners (9 F, 5 M) aged between 19 and 33 were recruited from the University of California, Berkeley community. Participants were native speakers of American English, without noticeable regional dialects. Participants reported normal hearing and no history of speech or language disorders. Listeners were paid for their participation.

2. Materials

Test stimuli were identical to those of Experiment 1 except that (1) speech/babble stimuli were processed as described above, and that (2) only the four place-of-articulation pairs /f/-/θ/, /v/-/ð/, /s/-/ʃ/, and /z/-/ʒ/ were tested, since these were the contrasts for which increased high-frequency content seemed to benefit normal-hearing listeners.

3. Procedures and apparatus

The procedure, task, presentation method, and adaptive procedure were identical to those of Experiment 1, except that since only four pairs were tested there was no break during the experiment. Testing took about 50 min.

4. Data analysis

As in Experiment 1, a repeated measure analysis of variance (ANOVA) with two within-subject factors (Style; 2 levels, Pair; 4 levels) and thresholds (dB snr) as dependent variable was performed. Acoustic measures and principal components were similarly compared across talkers with the clear speech intelligibility advantage.

C. Results and discussion

1. Fricative intelligibility for listeners with simulated hearing loss

Figure 3 shows snr thresholds as a function of pair type for clear and conversational fricative identification. For all place pairs except /f/-/θ/, clear speech showed lower snr thresholds relative to conversational speech. The Style \times Pair ANOVA showed an effect of Style [$F(1, 13)=13.9$, $p < .01$] with 2.5 dB lower thresholds for clear speech. There was also a Pair effect [$F(3, 39)=149.5$, $p < 0.001$], mostly derived from lower thresholds for sibilant pairs relative to non-sibilant pairs. The Style \times Pair interaction was also significant [$F(3, 39)=6.0$, $p < 0.01$]. Pairwise comparisons showed significant differences in thresholds as a function of style for /s/-/ʃ/ and /z/-/ʒ/ pairs, but not for non-sibilant pairs. In fact, for /f/-/θ/, clear speech resulted in higher (n.s.) thresholds compared to conversational speech. These results thus differed from Experiment 1 results in that (1) thresholds were on average much higher, (2) there was no clear speech effect for /f/-/θ/, and (3) the /z/-/ʒ/ pair showed the biggest clear speech effect, followed by /s/-/ʃ/, /v/-/ð/, and /f/-/θ/; in Experiment 1 the order was /z/-/ʒ/, /f/-/θ/, /s/-/ʃ/ and /v/-/ð/. On the other hand, the relative overall difficulty of fricative pairs was similar to Exp. 1; across speaking styles, the pair /s/-/ʃ/ resulted in the lowest thresholds, followed by /z/-/ʒ/, /v/-/ð/, and /f/-/θ/.

To determine how the loss simulation influenced the perception of fricatives in interaction with speaking style and contrastive pair, a three-way mixed model ANOVA was performed with two within-subject factors (Style, Pair) and listener group as a between-subject factor (two levels; Exp. 1 and Exp. 2). Since the four voicing distinction pairs were not included in Experiment 2, only the four place-of-articulation distinction pairs from Experiment 1 were considered. This analysis showed a main effect of Group [$F(1, 26)=26.4$, $p < 0.001$] with considerably (4.47 dB) higher thresholds for listeners with simulated hearing. A main effect of Style [$F(1, 26)=49.0$, $p < 0.001$] indicated, again, an overall clear speech advantage across listener groups. There was no Style \times Group interaction, suggesting that, on average, listeners with normal hearing and listeners with simulated impairment enjoyed comparable significant benefits from clear speech. The main-effect of Pair was significant [$F(3, 78)=212.8$, $p < 0.001$] but not the Pair \times Group interaction,

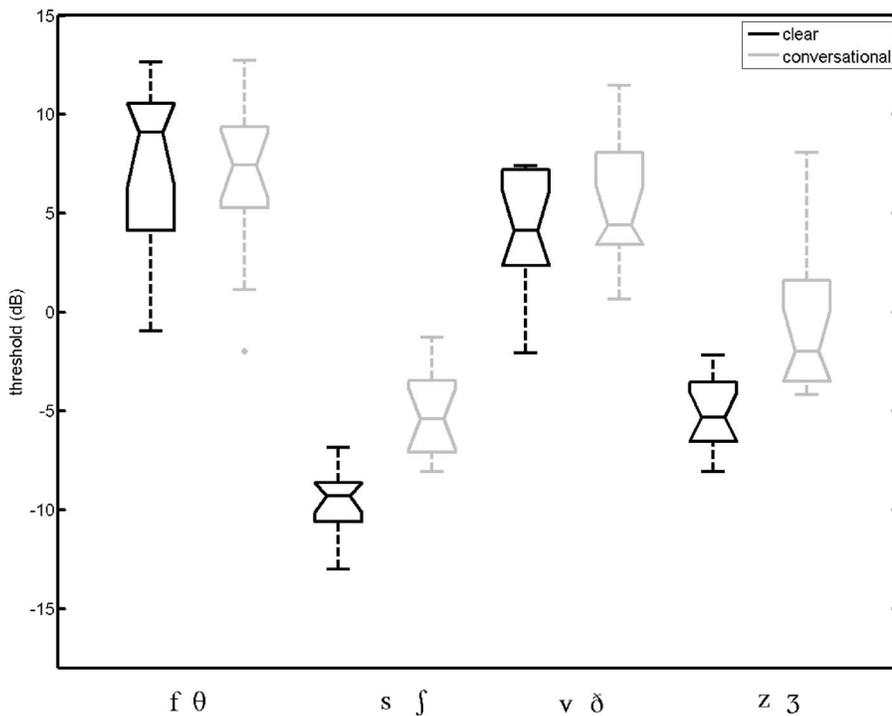


FIG. 3. Signal-to-noise ratio (snr) thresholds (dB) as a function of style and fricative pair in Experiment 2.

reflecting the common difficulty hierarchy mentioned above. Again, pairwise comparisons indicated that all four pairs were significantly different from each other, and that the effect was most notably derived from differences between sibilant and non-sibilant pairs. A $\text{Style} \times \text{Pair}$ interaction [$F(3,78)=212.8, p<0.01$] indicated that, across listener groups, the clear speech effect differed depending on the fricative pair. The $\text{Style} \times \text{Pair} \times \text{Group}$ interaction was significant [$F(3,78)=2.9, p<0.05$]; *post-hoc* tests suggested that the interaction was related to an increase in the magnitude of the clear effect for sibilants, and a decrease in the effect for non-sibilants, in the simulated impairment condition. This finding is illustrated in Fig. 4, which shows the clear speech effect as a function of pair and listening condition. It seems likely that, since non-sibilants are characterized by the highest peak and F2 values with a diffuse spread of energy below 10 kHz, important spectral cues for these

sounds are less audible/available to listeners with sloping hearing loss the higher they are transposed. Sibilants, on the other hand, have both higher relative amplitudes and more potential cues (esp. palatoalveolar peak frequencies) involving energy in lower regions. These cues would be better preserved in stimuli with simulated sloping loss.

2. Acoustic correlates of intelligibility benefit for listeners with simulated hearing impairment

In Experiment 2, individual talkers appeared on average in 168 (std. 12.4) clear and 168 (11.63) conversational trials. Again, averaged across listeners, contrasts, and snr values, the clear-minus-conversational difference in accuracy (% correct) varied considerably across speakers, from -6% to +18% (mean 3.9%, std. 6.6%). As discussed in Experiment 1, individual speakers' previously reported average style-

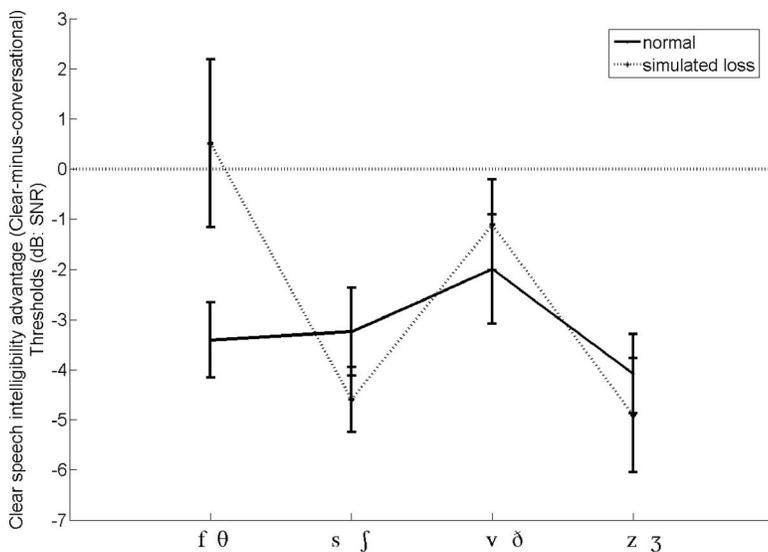


FIG. 4. Clear speech intelligibility advantage (clear-minus-conversational thresholds) in dB snr for listeners with normal hearing and listeners with simulated hearing impairment as a function of fricative pair.

related differences in production were compared with their style-related intelligibility differences in a first effort to relate clear speech benefits to specific acoustic modifications.

The results of individual measure correlations are shown in Table I. Overall, correlations were much less consistent than for Experiment 1; in particular, conspicuously absent were the positive correlations with spectral measures indicating shifts to higher frequency regions that were seen for place contrasts in Exp. 1. Since the perception of sibilant and non-sibilant pairs was affected differentially by the impairment, another set of correlation analyses compared intelligibility differences across speakers with acoustic differences separately for each class of sounds. While this comparison was considerably less well powered than the others described above, the results were potentially interesting and are also included in Table I. For sibilant pairs, *positive* correlations were seen between intelligibility advantages and spectral moment 3, and *negative* correlations with peak locations, at several window locations. For non-sibilant pairs, correlations were weaker and less straightforward. No significant correlations (all $p > 0.3$) were seen between clear speech advantages, either overall or considering sibilants and non-sibilants separately, with any of the acoustic principal components discussed above. Interestingly, the (nonsignificant) correlation between the first component (related to high-frequency energy) and the benefit for sibilants was negative.

IV. DISCUSSION

A. Overall clear fricative intelligibility

In two experiments, lower snr identification thresholds for place of articulation identification were seen for clear relative to conversational fricatives, indicating that, on average, clearly produced fricatives are more intelligible for both young normal-hearing listeners and listeners with simulated sloping, recruiting hearing impairment. In addition, clear speech was beneficial to normal-hearing listeners for voicing distinctions. However, these effects were not as uniform and robust across fricatives and listener groups as might have been expected. In Experiment 1, sibilant fricatives were easier to identify than non-sibilants for normal-hearing listeners overall, and clear speech provided slightly greater intelligibility benefits for sibilants than non-sibilants. Experiment 2 showed that these trends were exaggerated for simulated hearing-impaired listeners. In particular, a clear speech effect was seen *only* for sibilants, and clear speech may have even hurt intelligibility for voiceless non-sibilants, the worst-recognized sounds. These results are consistent with the notion (e.g., Ferguson and Kewley-Port, 2002) that the perceptual effects of clear speech acoustic modifications may be population dependent, and may interact in complex ways with different types of hearing impairment. As discussed below, they probably derive from differences in the audibility and weighting of acoustic cues across fricatives and listening conditions.

B. Acoustic and talker-related correlates of clear speech intelligibility effect

Comparison of individual speakers' estimated clear-speech intelligibility advantages with their previously reported (Maniwa *et al.*, submitted) clear-speech acoustic modifications revealed correlations that may be informative as to the acoustic sources of the "clear speech effect" in fricatives. Specifically, for place-of-articulation distinctions, strong positive correlations were found between acoustic and perceptual clear-vs-conversational differences for spectral measures, especially at central locations, including peak locations, M1, and spectral slope before peak locations. In addition, there were negative correlations between intelligibility improvement and increases in M3 and M4, and for intensity below 500 Hz. These results indicate that, overall, greater source strength (produced by higher volume velocity) in clear speech, resulting in spectral distributions with higher-frequency, more defined peaks, and more positive tilt, contributed to the intelligibility enhancement for place distinctions. Of course, it is more likely that these "global" changes in conjunction with higher-order patterns specific to individual fricatives and contexts actually led to the intelligibility effects that were seen. The experiments described here could not address this possibility, since within individual subtests snr values were not sufficiently equalized across speakers to make more specific comparisons. Probably for this reason, no strong correlations were seen relating acoustic measures and voicing intelligibility. In particular, acoustic results suggested that phonetic distance in terms of the voicing distinction was often "enhanced" in clear speech by increasing (or decreasing to a lesser degree) values for one class of fricatives while decreasing (or increasing to a lesser degree) values for the class. For example, intensity below 500 Hz decreased much less, and HNR significantly increased, for voiced fricatives whereas these values significantly decreased for voiceless fricatives. Similarly, noise duration and f_0 increased for both voiceless and voiced fricatives in clear speech, but to a much greater extent for voiceless fricatives. These differences in clear speech manipulations, and their perceptual effects, would have mostly been obscured by the analysis described here.

Our previous study (Maniwa *et al.*, submitted) also indicates that voiceless non-sibilants have, in addition to very low amplitudes, very high peak frequencies (higher than /s/), mean frequency, and F2, across speaking styles, and that these values are even higher in clear speech. This was probably a cause of the lack of clear speech benefits for (especially voiceless) non-sibilants, since the simulated impairment targeted higher frequencies (and low amplitudes). Sibilants, on the other hand, were characterized by more and lower energy, in some cases (esp. palato-alveolars) even more so in clear speech, so more potential cues for these sounds were preserved in the loss simulation. As a result of these differences, for listeners with simulated hearing impairment few overall correlations between acoustic and intelligibility differences in clear speech were apparent in Experiment 2. For identification of sibilant pairs specifically, contrary to Experiment 1 results, there were some *negative* correlations between acoustic changes in peak frequencies

and enhanced intelligibility (and marginal *positive* correlation between M3 and intelligibility advantage). This suggests that the lower the spectral information moved for palato-alveolar fricatives in clear speech, the more intelligible these sounds were, because this information was better preserved in the impairment simulation. Fewer and less consistent patterns could be seen to relate non-sibilant acoustic modifications to intelligibility. In other words, elevated thresholds and loudness recruitment influenced listeners' cue weighting for the perception of fricative sounds.

There were no Style \times Gender interactions in either experiment, indicating that female and male talkers did not differ in terms of the effectiveness of their clear speech acoustic modifications for intelligibility (cf. Bradlow *et al.*, 2003).

C. Conclusion

This study showed that clear speech enhanced the intelligibility of fricatives for both listeners with normal hearing and listeners with simulated hearing impairment. However, the effect was fricative and population dependent; notably, compared to normal-hearing listeners, impaired listeners showed reduced clear speech effects for non-sibilant place of articulation distinctions. Likewise, apparent acoustic correlates of the clear speech benefit differed across populations. For normal-hearing listeners, intelligibility benefits seemed to correlate with moves toward higher frequency regions for important cues; these patterns were generally not seen for impaired listeners, and may even have been reversed for some sounds. These results are straightforwardly explained based on audibility of cues at different levels and frequencies. We leave for future study a more thorough investigation of potential higher-order acoustic correlates of the clear speech effect in fricatives; this could be accomplished straightforwardly by using the results of the adaptive design described here to inform blocked-design experiments that are optimally controlled (and powered) for the distribution of fricatives, styles, and snr values across speakers and tokens. It will also be necessary to measure perception by actual hearing-impaired listeners in order to characterize the population-based differences we observed more quantitatively.

ACKNOWLEDGMENT

Portions of this research were conducted as part of K.M.'s doctoral dissertation under the supervision of A.J.

Behrens, S. J., and Blumstein, S. E. (1988). "On the role of the amplitude of the fricative noise in the perception of place of articulation in voiceless fricative consonants," *J. Acoust. Soc. Am.* **84**, 861–867.

Boothroyd, A. (1984). "Auditory perception of speech contrasts by subjects with sensorineural hearing loss," *J. Speech Hear. Res.* **27**, 134–144.

Bradlow, A. R., and Bent, T. (2002). "The clear speech effect for non-native listeners," *J. Acoust. Soc. Am.* **112**, 272–284.

Bradlow, A. R., Kraus, N., and Hayes, E. (2003). "Speaking clearly for children with learning disabilities: Sentence perception in noise," *J. Speech Lang. Hear. Res.* **46**, 80–97.

Chen, F. R. (1980). "Acoustic characteristics and intelligibility of clear and conversational speech at the segmental level," Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge, MA.

Cole, R. A., and Cooper, W. E. (1975). "Perception of voicing in English

affricates and fricatives," *J. Acoust. Soc. Am.* **58**, 1280–1287.

Dubno, J. R., and Levitt, H. (1981). "Predicting consonant confusions from acoustic analysis," *J. Acoust. Soc. Am.* **69**, 249–261.

Dubno, J. R., Dirks, D. D., and Langhofer, L. R. (1982). "Evaluation of hearing-impaired listeners using a Nonsense-syllable Test. II. Syllable recognition and consonant confusion patterns," *J. Speech Hear. Res.* **25**, 141–148.

Ferguson, S. H. (2004). "Talker differences in clear and conversational speech: Vowel intelligibility for normal-hearing listeners," *J. Acoust. Soc. Am.* **116**, 2365–2373.

Ferguson, S. H. (2002). "Vowels in clear and conversational speech: Talker differences in acoustic features and intelligibility for normal-hearing listeners," Doctoral dissertation, Indiana University, Bloomington, IN.

Ferguson, S. H., and Kewley-Port, D. (2002). "Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **112**, 259–271.

Fitzgibbons, P. J., and Wightman, F. L. (1982). "Gap detection in normal and hearing-impaired listeners," *J. Acoust. Soc. Am.* **72**, 761–765.

Gagné, J. P., Masterson, V. M., Munhall, K. G., Bilida, N., and Querengesser, C. (1994). "Across talker variability in auditory, visual, and audiovisual speech intelligibility for conversational and clear speech," *J. Acad. Rehabil. Audiol.* **27**, 135–158.

Gagné, J. P., Querengesser, C., Folkeard, P., Munhall, K. G., and Masterson, V. M. (1995). "Auditory, visual, and audiovisual speech intelligibility for sentence-length stimuli: An investigation of conversational and clear speech," *Volta Rev.* **97**, 33–51.

Gagné, J. P., Rochette, A.-J., and Charest, M. (2002). "Auditory, visual and audiovisual clear speech," *Speech Commun.* **37**, 213–230.

Glasberg, B. R., and Moore, B. C. J. (1992). "Effects of envelope fluctuations on gap detection," *Hear. Res.* **64**, 81–92.

Glasberg, B. R., and Moore, B. C. J. (1989). "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear impairments and their relationship to the ability to understand speech," *Scand. Audiol. Suppl.* **32**, 1–25.

Glasberg, B. R., Moore, B. C. J., and Bacon, S. P. (1987). "Gap detection and masking in hearing-impaired and normal-hearing subjects," *J. Acoust. Soc. Am.* **81**, 1546–1556.

Guerlekian, J. A. (1981). "Recognition of the Spanish fricatives /s/ and /f/," *J. Acoust. Soc. Am.* **70**, 1624–1627.

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

Hedrick, M. S. (1997). "Effect of acoustic cues on labeling fricatives and affricates," *J. Speech Lang. Hear. Res.* **40**, 925–938.

Hedrick, M. S., and Carney, A. E. (1997). "Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants," *J. Speech Lang. Hear. Res.* **40**, 1445–1457.

Hedrick, M. S., and Ohde, R. N. (1993). "Effect of relative amplitude of friction on perception of place of articulation," *J. Acoust. Soc. Am.* **94**, 2005–2026.

Hedrick, M. S., and Younger, M. S. (2003). "Labeling of /s/ and /ʃ/ by listeners with normal and impaired hearing, revisited," *J. Speech Lang. Hear. Res.* **46**, 636–648.

Heinz, J. M., and Stevens, K. N. (1961). "On the properties of voiceless fricative consonants," *J. Acoust. Soc. Am.* **33**, 589–596.

Helfer, K. (1997). "Auditory and auditory-visual perception of clear and conversational speech," *J. Speech Lang. Hear. Res.* **40**, 432–443.

Helfer, K. (1998). "Auditory and auditory-visual recognition of clear and conversational speech by older adults," *J. Am. Acad. Audiol.* **9**, 234–242.

Hughes, G. W., and Halle, M. (1956). "Spectral properties of fricative consonants," *J. Acoust. Soc. Am.* **28**, 303–310.

Iverson, P., and Bradlow, A. R. (2002). "The recognition of clear speech by adult cochlear implant users," in *Temporal Integration in the Perception of Speech*, edited by S. Hawkins and N. Nguyen (Cambridge: Center for Research in the Arts, Aix-en-Provence, France, Social Sciences, and Humanities), p. 78.

Jesus, L. M. T., and Shadle, C. H. (2002). "A parametric study of the spectral characteristics of European Portuguese fricatives," *J. Phonetics* **30**, 437–464.

Jongman, A. (1989). "Duration of friction noise required for identification of English fricatives," *J. Acoust. Soc. Am.* **85**, 1718–1725.

Jongman, A., Wang, Y., and Sereno, J. (2000). "Acoustic and perceptual properties of English fricatives," *Proceedings of the International Conference on Spoken Language Processing*, Beijing, China, **II**, 511–514.

King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., and Supowit,

- A. (1994). "Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation," *Vision Res.* **34**, 885–912.
- Krause, J. C., and Braida, L. D. (2002). "Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility," *J. Acoust. Soc. Am.* **112**, 2165–2172.
- Liu, S., Del Rio, E., Bradlow, A. R., and Zeng, F.-G. (2004). "Clear speech perception in acoustic and electric hearing," *J. Acoust. Soc. Am.* **116**, 2373–2383.
- Maniwa, M., Jongman, A., and Wade, T. (2007). "Acoustic and perceptual properties of clearly produced fricatives," Unpublished doctor's dissertation, The University of Kansas, Lawrence, KS.
- Mathworks, Inc., The. (2000). "MATLAB, The language of technical computing, version 7.0.0.19920."
- McCasland, G. P. (1979a). "Noise intensity and spectrum cues for spoken fricatives," *J. Acoust. Soc. Am.* **65**, S78–S79.
- McCasland, G. P. (1979b). "Noise intensity cues for spoken fricatives," *J. Acoust. Soc. Am.* **66**, S88.
- Miller, G. A., and Nicely, P. A. (1955). "An analysis of perceptual confusions among some English consonants," *J. Acoust. Soc. Am.* **27**, 338–352.
- Moore, B. C. J., and Glasberg, B. R. (1993). "Simulation of the effects of loudness recruitment and threshold elevation on the intelligibility of speech in quiet and in a background of speech," *J. Acoust. Soc. Am.* **94**, 2050–2062.
- Moore, B. C. J., and Peters, R. W. (1992). "Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity," *J. Acoust. Soc. Am.* **91**, 2881–2893.
- Nittrouer, S. (1992). "Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries," *J. Phonetics* **20**, 351–382.
- Nittrouer, S. (2002). "Learning to perceive speech: How fricative perception changes, and how it stays the same," *J. Acoust. Soc. Am.* **112**, 711–719.
- Nittrouer, S., and Miller, M. E. (1997a). "Predicting developmental shifts in perceptual weighting schemes," *J. Acoust. Soc. Am.* **101**, 2253–2266.
- Nittrouer, S., and Miller, M. E. (1997b). "Developmental weighting shifts for noise components of fricative-vowel syllables," *J. Acoust. Soc. Am.* **102**, 572–580.
- Owens, E. (1978). "Consonant errors and remediation in sensorineural hearing loss," *J. Speech Hear. Disord.* **43**, 331–347.
- Owens, E., Benedict, M., and Schubert, E. D. (1972). "Consonant phonemic errors associated with pure-tone configurations and certain kinds of hearing impairments," *J. Speech Hear. Res.* **15**, 308–322.
- Payton, K. L., Uchanski, R. M., and Braida, L. D. (1994). "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," *J. Acoust. Soc. Am.* **95**, 1581–1592.
- Picheny, M. A., Durlach, N. I., and Braida, L. D. (1985). "Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech," *J. Speech Hear. Res.* **28**, 96–103.
- Raphael, L. (1972). "Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English," *J. Acoust. Soc. Am.* **51**, 1296–1303.
- Schum, D. (1996). "Intelligibility of clear and conversational speech of young and elderly talkers," *J. Am. Acad. Audiol.* **7**, 212–218.
- Sher, A. E., and Owens, E. (1974). "Consonant confusions associated with hearing loss above 2000 Hz," *J. Speech Hear. Res.* **17**, 669–681.
- Singh, S., and Black, J. W. (1966). "Study of twenty-six intervocalic consonants as spoken and recognized by four language groups," *J. Acoust. Soc. Am.* **39**, 372–387.
- Slaney, M. (1998). "Auditory Toolbox, version 2.0," <http://cobweb.ecn.purdue.edu/~malcolm/interval/1998-010/>. last viewed 25 September 2007.
- Soli, S. D. (1982). "Structure and duration of vowels together specify fricative voicing," *J. Acoust. Soc. Am.* **72**, 366–378.
- Soli, S. D., and Arabie, P. (1979). "Auditory versus phonetic accounts of observed confusions between consonant phonemes," *J. Acoust. Soc. Am.* **66**, 46–58.
- Stevens, K. N. (1985). "Evidence for the role of acoustic boundaries in the perception of speech sounds," in *Phonetic Linguistics: Essays in honor of Peter Ladefoged*, edited by V. Fromkin (Academic, New York), pp. 243–255.
- Stevens, K. N., Blumstein, S. E., Glicksman, L., Burton, M., and Kurowski, K. (1992). "Acoustic and perceptual characteristics of voicing in fricatives and fricative clusters," *J. Acoust. Soc. Am.* **91**, 2979–3000.
- Uchanski, R. S., Choi, S. S., Braida, L. D., Reed, C. M., and Durlach, N. I. (1996). "Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate," *J. Speech Hear. Res.* **39**, 494–509.
- Villchur, E. (1974). "Simulation of the effect of recruitment on loudness relationships in speech," *J. Acoust. Soc. Am.* **56**, 1601–1611.
- Wang, M. D., and Bilger, R. C. (1973). "Consonant confusions in noise: A study of perceptual features," *J. Acoust. Soc. Am.* **54**, 1248–1266.
- Whalen, D. H. (1981). "Effects of vocalic formant transitions and vowel quality on the English [s]-[ʃ] boundary," *J. Acoust. Soc. Am.* **69**, 275–282.
- Zeng, F.-G., and Turner, C. W. (1990). "Recognition of voiceless fricatives by normal and hearing-impaired subjects," *J. Speech Hear. Res.* **33**, 440–449.