

SPEECH CUE ENHANCEMENT FOR THE HEARING IMPAIRED:
 III. AMPLIFICATION OF FRICATION FOR IMPROVED PERCEPTION OF FINAL FRICATIVE VOICING

REVOILE, S.G., HOLDEN-PITT, L., EDWARD, D., and PICKETT, J.M.

Center for Auditory and Speech Sciences
 Gallaudet Research Institute
 Gallaudet University

ABSTRACT

For 20 severely/profoundly hearing-impaired listeners, voicing perception for final fricatives was tested using spoken syllables with and without enhanced frication. The enhancements involved filtering and amplification of /f/, /s/, and, of iterations of pitch periods in the vowel offset to replace /v/, /z/. Nearly 3/4 of the listeners showed considerable improvement in perception for the fricatives with enhancement. Discrimination training for voicing cognate syllables contributed to the improvement in perception found for the enhanced fricatives. For syllables with natural fricatives, lower performance and less improvement was seen as a result of training. Inaudibility of the enhanced fricatives seemed to explain the poor performances of listeners who showed no improvement in voicing perception for the enhanced fricatives.

INTRODUCTION

Previously, we examined the effects of cue degradation on final consonant voicing perception by moderately/severely hearing-impaired listeners (Revoile et al., 1982^[1] and 1985^[2]). Elimination of the vowel duration cue and/or the consonant constriction cues (i.e., frication, stop bursts, presence of voiced murmur) degraded perception for most of the hearing-impaired listeners studied. We then began to investigate whether enhancement or exaggeration of cues to final consonant voicing could yield improved perception for hearing-impaired persons who typically manifest reduced ability to distinguish consonant voicing (Revoile et al., 1987^[3], 1986^[4] & 1986^[5]). This paper describes an experiment on amplification and filtering of final fricative consonants as a means of facilitating fricative voicing distinctions by hearing-impaired listeners.

METHOD

The listeners were 20 hearing-impaired Gallaudet undergraduates who had predominantly profound losses. Their threshold averages for .5, 1, and 2 kHz (3FA) ranged from 80 to 108 dB HL, with a mean (\bar{X}) of 94 dB HL. The listeners were selected to have 3FA of at least 80 dB HL because our previous findings suggested that

reduced consonant voicing perception would be concomitant with such impairments (Revoile et al., 1982^[1], 1985^[2]).
Stimuli. The syllables /bæ s/, bæ z, bæ f, bæ v/ served as the stimuli to test fricative voicing perception. For most of the experiment, 4 different syllable utterances (female talker) representing each fricative were used in a test block (16 different utterances). These tokens had been selected from a larger pool of utterances to differ minimally in vowel duration between voicing cognate syllables. Some mean acoustical characteristics of the test utterances are shown in Table 1.

The phoneme segment durations were measured visually on waveforms of the utterances; rms intensity was measured for the duration of each segment. The vowel/frication boundary was identified on the waveform as the point at which periodic vowel oscillation ceased. The utterances had been digitized (16.67 kHz sampling rate) for these measurements and for processing of the frication enhancements. Further details of the recording and measurement procedures can be found in Revoile et al. (1986^[4]).

The frication segments of each utterance were altered to generate the enhanced stimuli. The /f/ and /s/ frications were low-pass filtered (5 kHz) and then amplified by 21 dB. The /v/ and /z/ frications were deleted and replaced by 4 iterations of 2 to 4 pitch periods copied from the end of the vowel in each utterance. These pitch periods contained some f_0 as well as consonant friction noise. The pitch periods were iterated to yield frication durations that nominally matched those of the deleted frications. The segments were band-pass filtered (.25 to 1 kHz) to reduce the presence of the vowel f_0 and intensified by 18 dB. In the enhanced syllables (Table 1), note that the frications had been amplified to approximate the level of the vowels.

Procedures. The testing and training of fricative voicing perception was carried out in an experiment on cue enhancement for final consonant voicing that lasted about 2-1/2 months. During that period, the listeners participated in twice weekly sessions of 50 minutes each. The initial 3 to 4 sessions of the experiment were devoted to screening tests for perception of final consonant voicing, to insure that the students chosen as listeners manifested reduced voicing perception of final fricatives. The training segment of the experiment began with a

TABLE 1. A summary of some acoustical characteristics from 4 utterances each of /bæ f/, /bæ s/, /bæ v/, /bæ z/.

		Duration, ms		Intensity, dB ¹		
		Vowel	Frication	Vowel	Frication	
					Unaltered	Enhanced ²
/bæ f/	\bar{X}	289.7	211.3	49.1	27.9	47.8
&	S.D.	(8.4)	(18.5)	(1.2)	(2.4)	(0.4)
/bæ s/						
/bæ v/	\bar{X}	291.6	114.0	49.4	28.2	49.9
&	S.D.	(9.9)	(14.2)	(1.1)	(1.7)	(1.5)
/bæ z/						

¹ Relative to an arbitrary reference.

² Intensity measurements for the enhanced voiced fricatives were of iterations of pitch periods from the vowel offset, which were substituted for the natural /v/ and /z/ frications.

baseline assessment of the syllables both unaltered and with frication enhancement. Training for the syllables with enhanced frications was usually followed by training for the unaltered syllables. All stimuli were presented to each listener's better ear at a most comfortable level (MCL) established by an adaptive procedure at the beginning of each session. Syllable identification trials were used to test fricative voicing perception throughout the experiment. The response buttons used by the listeners were labeled: BAFF, BASS, BAV, BAZZ, and BA. No feedback of correct answers was given.

The unaltered and enhanced utterances were tested in separate blocks of trials. In each block of 48 trials, the 4 different utterances per fricative were each presented 3 times in random order. These syllable blocks were tested for the baseline measurements of fricative voicing perception, and before and after training.

The training of fricative voicing perception involved discrimination trials of voicing cognate syllables with feedback following each trial. In separate blocks, two types of discrimination trials were used--an "oddy" procedure and a "paired-comparison" procedure. The "oddy" procedure presented syllable sequences, such as BASS BAZZ BASS or BAZZ BASS BAZZ in separate trials. The listener chose which sequence had been presented. The "paired-comparison" procedure presented syllables pairs such as BASS BASS, BAZZ BASS, BAZZ BAZZ, or BASS BAZZ and the listener selected the pair heard.

The unaltered versus enhanced syllables were used in separate training sessions. A training session began with assessment of syllable identification. Subsequently the "oddy" followed by "paired comparison" discrimination procedures were presented for each of the syllable sets representing /s/-/z/ and /f/-/v/. Finally, syllable identification was retested at the end of the session.

Most of the the listeners were administered 2 to 6 training sessions (\bar{X} =3.8) for the syllables with enhanced frications. Typically, listeners who received fewer training sessions were those who evidenced chance performance after training for the fricative-enhanced utterances. Because of their poor performances, these listeners were not trained for the unaltered syllables, to limit their frustration with the training. However, all listeners who showed at least some ability to perceive voicing after training for the enhanced frications (>60%) received 1 to 3 training sessions (\bar{X} =1.7) for the unaltered syllables.

An additional discrimination procedure, "frication presence", was used to determine whether the frications were audible to the listeners. The procedure required discrimination of an utterance-with-frication, versus that same utterance with frication deleted. The trials used were 3IFC, with the fricationless utterance presented twice and the full utterance once. The listener chose which interval contained the full utterance. Feedback of the correct answer was presented after each trial.

RESULTS

The performance of each listener for each fricative was scored according to percent correct voicing perception (errors in place perception ignored) for each block of syllables presented. A total percent correct score was computed for each test condition (i.e., screening, baseline, etc.) by averaging scores across frications and repetitions of stimulus blocks. Much of the data represented chance level performance. When such performance was seen among the listeners overall, no statistical analyses were carried out.

The listeners were assigned to one of 2 groups depending upon whether their average performances were above chance level (at least 60%) for voicing perception on the tests administered after training for the enhanced frications. The

6 listeners who scored below this criterion formed the poor group; their mean (X) 3FA was 98 dB HL. The other 14 listeners (X 3FA = 91 dB HL) composed the good group. Mean fricative voicing perception for the groups on each test condition is summarized in Table 2.

Performance for unaltered syllables. The good and poor listener groups can be compared for performance on the unaltered syllables when these stimuli were administered for the screening and for the baseline tests. For both of these tests, similar performances occurred between the groups; mean scores of less than 50% were obtained. These chance level performances indicate that the voicing cues for the natural fricatives were generally not utilized by the listeners, at least prior to training. (Recall that the vowel duration cue had been reduced by pre-selection of voicing cognate utterances having minimal value for this cue.)

For the good group, training for the unaltered syllables resulted in some improvement in voicing perception for the fricatives. The group's scores after training for the unaltered syllables were significantly higher than before [t-test for paired samples: $t(13)=3.9$, $p<.01$], although by a mean difference of just 8%. Because only 1 to 3 sessions of training were used for the unaltered syllables, it's possible that the maximum improvement in fricative voicing perception was not achieved. Lack of familiarity with the training procedure was probably not a factor in limiting the improvement seen for the unaltered syllables because their training was subsequent to that for the enhanced syllables. Nevertheless, a longer period of training might have heightened the listeners' awareness of the natural fricative voicing cues in the unaltered syllables.

Performance for syllables with frication enhancement. When the syllables with enhanced fricatives were initially presented, neither group of listeners performed well. The baseline

tests for these syllables, administered just before training, yielded performance at chance level for both groups.

The training for the fricative-enhanced syllables resulted in large performance differences between the listener groups, both before and after training. The poor group performed at chance level for the syllables with enhanced fricatives, regardless of training, while the good group scored at least 20% above chance. An analysis of variance was carried out using as factors: group (good versus poor) by before/after (treated as a repeated measure). Highly significant differences were found between the good and poor groups for performance with the fricative-enhanced syllables during training [$F(1,18)=45.9$, $p<.001$].

An interaction appeared between the factors: group and before/after [$F(1,18)=5.6$, $p=.03$], indicating that the amount of performance change occurring before versus after training was different between the listener groups. Further analyses of simple main effects revealed that the scores after training were significantly greater than before training for the good group [$F(1,18)=27.7$, $p<.001$], while the poor group showed no difference in scores before versus after training [$F(1,18)=.39$, $p=.54$].

The results for the discrimination of frication presence were examined to determine whether the groups' performance differences for the enhanced syllables might be associated with the general audibility of the enhanced frications. Between utterances with enhanced frication versus those utterances fricationless, the good group showed 87% mean discrimination and the poor group, 51%. While the poor group scored somewhat above chance (33%), their reduced level of performance indicates that they were unable to discriminate the presence versus absence of the enhanced frications for the majority of the utterances. This reveals that these listeners' inability to perceive voicing for the enhanced

fricatives was generally a result of frication inaudibility. For the unaltered syllables, the mean discrimination score for presence-vs-absence of the natural frications was 51% for the good group and 28% for the poor group.

Unaltered versus enhanced training effects. For the good group of listeners, scores could be compared between the unaltered and enhanced syllables for training of fricative voicing perception. Overall, performance with the enhanced syllables throughout training was markedly better than for the unaltered syllables [$F(1,13)=61.2$, $p<.001$]. However, the effects of training were similar for the unaltered and enhanced syllables as indicated by the absence of interaction for scores representing the two types of syllables before and after training [$F(1,13)=.31$, $p=.59$].

It was of interest to examine whether the level of perceptibility differed between /f/, /s/ versus /v/, /z/, especially because of the disparate spectral characteristics of enhanced /v/, /z/. However, this seemed to have no effect on the listeners voicing perception for the fricatives, as scores for /f/, /s/ versus /v/, /z/ were similar [$F(1,13)=1.4$, $p=.25$].

DISCUSSION

This study revealed that frication filtering and amplification yielded improved perception of voicing for syllable-final fricatives by almost 3/4 of the severely/profoundly hearing-impaired subjects tested. The improvement in perception occurred for enhanced fricatives that were amplified to the level of the preceding vowels in /bæC/ syllables. When the frications were at natural levels in the syllables, i.e., -21 dB re the vowels, the listeners' performances were somewhat above chance level, but only after discrimination training for voicing cognate syllables. For the fricative-enhanced syllables, the discrimination training facilitated the listeners' improved perception of fricative voicing.

The remaining listeners, about 1/4 of the total group, showed no improvement in voicing perception for the fricative-enhanced syllables, even after training. For these listeners, most of the enhanced frications were probably inaudible. A discrimination test for presence/absence of the frications in the test syllables revealed that these listeners were unable to distinguish most of the test utterances when presented with versus without enhanced frications. Further evidence of the inaudibility of the enhanced fricatives is apparent from the syllable presentation levels used relative to the tone sensitivity of these listeners. Based on the listening levels chosen by these subjects (\bar{X} MCL = 113 dB SPL), the vowels were presented at

sensation levels (SL) of less than 15 dB. The SLs of the enhanced frications were probably lower, due to the listeners' greater hearing loss in the frequency regions of the frication spectra compared to the vowel spectra.

The amount of improvement in final consonant voicing perception effected by the enhanced fricatives in this study is somewhat less than that seen for amplified stop bursts (Revoile, et al., 1987 [3]) and enhancement of the vowel duration cue (Revoile, et al., 1986 [4]). In future studies, we will investigate the relative effects of single versus multiple cue enhancements in search of the maximum improvement for consonant recognition by hearing-impaired listeners.

CONCLUSIONS

Some severely/profoundly hearing-impaired listeners performing at chance level for fricative voicing perception may obtain improved performance as a result of frication filtering and amplification to a level comparable to that of the preceding vowel. Discrimination training for syllables with enhanced voiceless versus voiced fricatives seems important in effecting the improved perception of fricative voicing.

REFERENCES

- [1] Revoile, S., Pickett, J. M., Holden, L.D., and Talkin, D. (1982). Acoustic cues to final stop voicing for impaired-and normal-hearing listeners. *J. Acoust. Soc. Am.* 72, 1145-1154.
- [2] Revoile, S., Holden-Pitt, L., and Pickett, J. M. (1985). Perceptual cues to the voiced-voiceless distinction of final fricatives for listeners with impaired or normal hearing. *J. Acoust. Soc. Am.*, 77, 1263-1265.
- [3] Revoile, S., Pickett, J., Holden-Pitt, L., Edward, D., Pickett, J. and Brandt, F. (1987, in review). Speech-cue enhancement for the hearing-impaired: II. Amplification of burst/murmur cues for improved perception of final stop voicing. *J. Rehab. Res. Dev.*
- [4] Revoile, S., Holden-Pitt, L., Pickett, J., and Brandt, F. (1986) Speech cue enhancement for the hearing impaired: I. Altered vowel durations for perception of final fricative voicing. *J. Speech Hear. Res.*, 29, 240-255.
- [5] Revoile, S., Holden-Pitt, L., Edward, D., and Pickett, J. (1986) Some rehabilitative considerations for future speech-processing hearing aids. *J. Rehab. Res. Dev.*, 23, 89-94.

TABLE 2. For unaltered and enhanced final /f/, /s/, /v/, /z/ in /b C/ syllables, mean percent correct voicing perception by two groups of hearing-impaired listeners for various tests during the experiment.

		Unaltered Syllables				Syllables with Enhanced fricatives			
		Screening Tests ^a	Baseline	Training		Baseline	Training		
				Before	After		Before	After	
Good Group (n=14)	X (S.D.)	46.0 (6.2)	47.0 (8.5)	54.1 (7.7)	62.3 (8.7)	50.0 (15.1)	73.0 (10.9)	82.6 (8.0)	
Poor Group (n=6)	X (S.D.)	45.5 (9.9)	46.0 (11.6)			54.0 (7.9)	50.2 (6.9)	52.0 (5.1)	

^a Scores based on 40-utterance block. Dashes are inserted where no tests were administered.