Journal of Speech and Hearing Research, Volume 33, 440–449, September 1990

RECOGNITION OF VOICELESS FRICATIVES BY NORMAL AND HEARING-IMPAIRED SUBJECTS

FAN-GANG ZENG CHRISTOPHER W. TURNER
Communication Sciences and Disorders, Syracuse University

The purpose of this study was to investigate the sufficient perceptual cues used in the recognition of four voiceless fricative consonants [s, f, θ, ʃ] followed by the same vowel [iː] in normal-hearing and hearing-impaired adult listeners. Subjects identified the four CV speech tokens in a closed-set response task across a range of presentation levels. Fricative syllables were either produced by a human speaker in the natural stimulus set, or generated by a computer program in the synthetic stimulus set. By comparing conditions in which the subjects were presented with equivalent degrees of audibility for individual fricatives, it was possible to isolate the factor of lack of audibility from that of loss of suprathreshold discriminability. Results indicate that (a) the frication burst portion may serve as a sufficient cue for correct recognition of voiceless fricatives by normal-hearing subjects, whereas the more intense CV transition portion, though it may not be necessary, can also assist these subjects to distinguish place information, particularly at low presentation levels; (b) hearing-impaired subjects achieved close-to-normal recognition performance when given equivalent degrees of audibility of the frication cue, but they obtained poorer-than-normal performance if only given equivalent degrees of audibility of the transition cue; (c) the difficulty that hearing-impaired subjects have in perceiving fricatives under normal circumstances may be due to two factors: the lack of audibility of the frication cue and the loss of discriminability of the transition cue.

KEY WORDS: speech perception, fricative, hearing loss

The main body of today's knowledge on fricative recognition is found in the pioneering studies of the acoustic analysis of fricatives conducted about 20–30 years ago (e.g., Fant, 1960; Fletcher, 1953; Harris, 1958; Heinz & Stevens, 1961; Hughes & Halle, 1956; Stevens, 1971; Strevens, 1960). These studies indicated that there are two kinds of acoustic cues that are most likely responsible for voiceless fricative recognition. One is the essentially steady-state spectral shape differences existing in the voiceless frication noise burst: [s] typically has a spectral peak at 4-8 kHz and [ʃ] has one at 2-4 kHz, whereas [f] and [θ] have similar flat spectra. The other cue is the dynamic spectral changes occurring during the voiced formant transition between the fricative and adjacent vowel, which particularly play a role in differentiating between [f] and [θ]. The perceptual cues for the voiced fricatives are presumed to have a similar explanation, except that there is an additional periodic vibration source at the glottis during most if not all of the frication portion of the syllable.

Other possible acoustic cues such as the duration and amplitude differences have been shown to have minimal perceptual significance. Jongman (1989) reported that shortening the frication noise to a duration of 70 ms does not significantly decrease fricative recognition performance compared to conditions where the full-duration cue is presented. The overall amplitude of the frication noise relative to the vowel also has been reported to be of some importance in distinguishing between different voiceless fricatives (Guerlekian, 1981; McCasland, 1979a, 1979b; and Stevens, 1985). But a recent study (Behrens & Blumstein, 1988) showed that the amplitude effect is small under more natural and controlled conditions where the spectral properties of the friction noise and the formant transitions are synthesized to be compatible.

Although specific investigations of fricative recognition by hearing-impaired subjects are rather rare compared with those of stop consonants, several studies considering the general issue of consonant recognition have shown that hearing-impaired subjects often have difficulty recognizing fricatives (Bilger & Wang, 1976; Owens, 1978; Owens, Benedict, & Schubert, 1972). The precise reasons for this difficulty are not clear, but several factors such as reduced audibility, poor frequency selectivity and impaired temporal resolution have been suggested (Van Tasell, 1981). In our view, the factors affecting hearing-impaired subjects' speech recognition can be divided into two basic categories: audibility and suprathreshold discriminability. Audibility refers to whether a specific speech cue is presented at a suprathreshold level to a hearing-impaired subject. Given that hearing-impaired listeners, by definition, have elevated thresholds, this would obviously be a major factor. However, it is also known that hearing-impaired subjects are not always capable of achieving normal recognition performance in cases where the speech cues are presented at audible levels (e.g., Turner & Robb, 1987). Poor frequency selectivity, impaired temporal resolution and abnormal loudness are possible factors that may cause a loss of suprathreshold discriminability.

These two factors—a lack of audibility and a loss of discriminability—have often been confused in previous experiments, especially when hearing-impaired subjects are involved. In order to avoid this potential confusion, the fricative stimuli will be presented to both normal-hearing and hearing-impaired subjects at an equivalent degree of audibility in the current investigation. In this case, fricative recognition can be compared on a common basis, in which the factor of loss of discriminability of a cue can be investigated separately from its lack of audi-

  0022-4685/90/3303-0440$01.00/0

bility. The purpose of this study is to answer the following questions:

1. What acoustic cues are sufficient for correct fricative recognition in normal-hearing subjects?
2. Are these cues audible (or presented at intensity levels above a subject's sensitivity thresholds) to hearing-impaired subjects under normal circumstances? If not, does the lack of audibility explain their errors in fricative recognition?
3. Given an equivalent degree of audibility of these cues to both normal-hearing and hearing-impaired subjects, do they achieve equivalent recognition performance, and if not, what conclusions regarding the abilities of hearing-impaired subjects can we draw from this comparison?

This paper consists of two main experiments. Experiment 1 investigates the sufficient cues for fricative recognition by normal-hearing subjects, in which both natural and synthetic speech stimuli are used. Experiment 2, on the other hand, attempts to answer questions 2 and 3 by comparing fricative recognition between normal-hearing and hearing-impaired subjects.

EXPERIMENT 1

There were two reasons for using both natural and synthetic stimuli in this experiment. We wanted to make sure that the results obtained from synthetic stimuli would be comparable to those with natural speech. On the other hand, because the interpretation of closed-set recognition results with natural speech may be complicated by variations that would normally be perceptually insignificant (e.g., amplitude and duration differences) it is often more convenient and straightforward to use the much more easily controlled synthetic speech.

METHOD

Subjects

Four graduate students, 3 female and 1 male, aged 22–25, served as the normal-hearing subjects. Their average pure-tone thresholds, determined by a four-alternative-forced-choice (4AFC) procedure (using a 2-dB step size and a TDH-39 earphone) tracking the 71%-correct level of detection (Levitt, 1971), are shown by the solid line with cross symbols in Figures 2 and 4.

Stimuli

Four fricative-vowel syllables—see, fee, thee, and she ([s, f, θ, ʃ] followed by the vowel [iː]) were chosen as the speech stimulus set, because they are simple and commonly used fricative and vowel combinations in conversational English. Three stimulus conditions were used in this experiment.

Condition 1. The natural stimulus set, produced by a 37-year-old, male, native English speaker, was digitized at a sampling frequency of 10 kHz. To avoid inclusion of undesired cues that are difficult to control during speaking, such as different durations or intensities, we performed digital duration and intensity normalization for these four natural speech tokens. For duration normalization, the Voicing-Onset-Time (VOT) for each token was identified based upon visual inspection of the waveform, which then served as a common mark for duration normalization. Second, each fricative-vowel syllable was limited to the 180 ms preceding the VOT (i.e., frication portion) and the 220 ms following the VOT (voicing portion); the remainder of the CV stimulus outside this 400-ms overall duration was eliminated. In addition, each stimulus was given 10-ms rise- and fall-times to avoid onset and offset clicks. For intensity normalization, we measured the maximum amplitude for all four CV stimuli; the amplitude of each stimulus was normalized to have this same maximum amplitude. Because the vowel has a much greater intensity than the fricatives, intensity normalization performed in this way can also be considered as vowel intensity normalization.

Condition 2. The same four fricative-vowel syllables were synthesized on an LSI-11/23 computer according to the parallel resonance procedure developed by Klatt (1980). The spectral characteristics of the frication portion of the fricatives were determined by the frequency, bandwidth, and amplitude settings of their formants. Table 1 gives the parameters chosen in synthesizing the four fricatives.

The frication portion for all four speech tokens was 200 ms in duration. The voicing started at 200 ms (at the offset of frication) and immediately following voicing onset there was a 40-ms formant transition period followed by a 160-ms vowel [iː]. Thus, the total duration for each of

TABLE 1. Parameter values for the synthesis of fricative consonants. The unit for the formant frequency (F) and its bandwidth (B) is Hz, the unit for the amplitude (A) is relative dB; A1 and A2 for the frication portion of all tokens are equal to zero; F4 = 3300, F5 = 3700, F6 = 4900 Hz.

Fricative	*F1*	*F2*	*F3*	*B1*	*B2*	*B3*	*A3*	*A4*	*A5*	*A6*	*AB*
[s]	320	1390	2530	200	80	200	0	0	0	66	0
[f]	340	1100	2080	200	120	150	0	0	0	0	50
[θ]	320	1290	2540	200	90	200	0	0	0	35	45
[ʃ]	300	1840	2750	200	100	300	57	48	48	50	0

those four fricative-vowel syllables was exactly 400 ms. Additionally, the fundamental frequency in the synthesized vowel decreased from 120 Hz to 90 Hz between the onset of voicing (at 200 ms) and the end of the stimulus (at 400 ms) to more closely simulate natural speech.

Condition 3. This set of stimuli was used to investigate the independent contribution from the frication cue to correct fricative recognition. By truncating the synthetic fricative-vowel tokens (in Condition 2) after 200 ms duration, only the initial or frication portion (duration = 200 ms) of the CV syllables was preserved. That is, all four stimuli in this condition contained only the steady-state frication portion and did not include any voicing or transition information. The truncated frication portion was also given a 10- ms fall time to avoid clicks. However, the nominal presentation level of these stimuli was defined as the rms level of the absent vowel portion, in order to facilitate comparison between experiments.

Digitized forms of both the natural and synthetic tokens were stored on an LSI-11/23 computer disk and presented to the subject's headphone through a low-distortion 12-bit digital-to-analog converter (Data Translation 3366) followed by a low-pass filter. The sampling rate was 10 kHz and the filter (Kemo-type VBF8) had a cutoff frequency at 5 kHz. The headphone (Telephonics TDH-39 mounted in an MX41/AR cushion) was shown to have a nearly flat frequency response up to 5 kHz, when measured in an NBS 9-A coupler. All sound levels in this study, whether for natural speech stimuli, pure-tone test stimuli, or synthetic speech stimuli, were measured in this NBS 9-A coupler.

Procedure

Subjects were seated individually in a sound-proof booth, in front of a response box with four buttons labeled as *see, fee, thee* and *she*. The subjects were instructed to press the button corresponding to the speech token presented and to guess if they were uncertain of the correct response. A 200-ms warning light preceded the onset of each stimulus, and a 600-ms delay followed each response. The computer was programmed to wait until the subject responded before initiating the next trial. The 600-ms period, together with the subject's response delay, served as the interstimulus interval. A single experimental block consisted of the four fricative-vowel speech tokens, each presented 45 times in randomized order, for a total of 180 trials. Each subject practiced for 2–3 hours with the test stimuli presented at a comfortable listening level before the actual data collection. The subjects were informed of the overall percent-correct score after each experimental block. Other than that, no feedback was given.

The speech recognition task was conducted across a wide range of presentation levels, from as low as 20 dB SPL for normal-hearing subjects to as high as 120 dB SPL for hearing-impaired subjects. This nominal presentation level was expressed as the rms level of the vowel portion of either the synthetic or the natural CV syllable (specifically, the vowel isolated from see was used as the reference for natural speech).

RESULTS

Analysis of Audibility Spectrum

Many investigations have indicated that listeners can sum energy across a critical bandwidth, or the bandwidth of an auditory filter, for the detection of narrow- and broad-band signals (Gässler, 1954; Green, 1958; Spiegel, 1979). An "audibility spectrum" for a broadband signal such as speech can be then defined in this study as the output of these auditory filters that sum, for each point on the frequency domain, the stimulus energy within a critical bandwidth centered at its frequency. The specific determination of an "audibility spectrum" for each speech stimulus in this study was similar to the procedure described in Turner and Robb (1987) in their study of stop consonants. The auditory filters used to calculate the audibility spectrum had an equivalent rectangular bandwidth equal to .16 times their center frequency. This meant that our audibility spectrum was similar to, but more detailed in shape than, a one-third octave band spectral representation. The validity of the audibility spectrum concept for predicting the audibility of our stimuli was further verified by calculating the audibility spectrum of single pure tones and comparing them with a subject's sensitivity threshold curve. In this case, the audibility spectrum of a pure tone showed a peak that was located at a level essentially equal to the subject's pure tone threshold at that frequency. Thus, our audibility spectrum serves as a proper comparison to an individual subject's pure tone thresholds, for the purpose of determining the audibility of spectral regions of our stimuli.

The audibility spectrum for the frication portions of the four natural tokens was calculated from the 50-101.2 ms voiceless interval of the normalized CV stimuli, whereas that for the transition portions was calculated from the 170-221.2 ms interval (the VOT is at 180 ms in each stimulus). It is especially informative to consider the differences in intensity between the various portions of the syllables. In order to express the average difference in intensity between frication, transition, and vowel for the four fricatives from this speaker, mean frication and transition audibility spectra were created by averaging the four fricatives' audibility spectra together. Figure 1 shows the average difference in intensity between the mean spectra of frication, transition, and vowel at a nominal presentation level of 70 dB SPL. It can be noted that the frication intensity is about 20–30 dB less than the vowel [iː] while the transition is about 5–10 dB less. It should be noted the vowel [iː] audibility spectrum was calculated from 280–331.2-ms interval of the normalized CV stimuli.

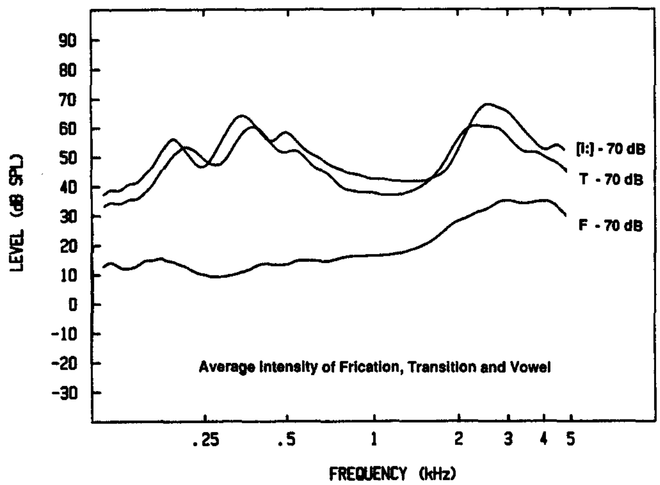

FIGURE 1. Comparison of the mean audibility spectra between the frication portion (F), transition portion (T) and vowel portion ([i:]) of the natural speech stimuli.

Recognition Results—Condition 1

Figure 2 shows average percent-correct recognition by the 4 normal-hearing subjects. Their mean pure-tone sensitivity thresholds are plotted along with the mean frication (upper panel) and transition (lower panel) audibility spectra at several nominal presentation levels. Substantial recognition of these natural fricative tokens occurred at presentation levels (20 and 30 dB SPL) where the frication portion of the stimuli was inaudible, while at least part of the transition portion was at suprathreshold levels. We consider this occurrence, related to the intensity differences between the various portions of the fricative syllables, to be an important finding and will further explore it throughout the remainder of the paper.

Results, as shown in Figure 2, were condensed into a more manageable form: a performance-audibility (PA) function that illustrates the relationship between the recognition performance at any given presentation level and a single-number measure of audibility for the CV tokens at that level (Turner & Robb, 1987). The single-number measure of audibility was defined as the percentage of logarithmically spaced points on the audibility spectrum that were above the subject's thresholds within a defined frequency range at a given presentation level.

The choice of the defined frequency range over which the percentage of audible spectrum was calculated was based upon pilot work performed in our laboratory and also upon acoustic analyses available in the literature (e.g., Minifie, Hixon, & Williams, 1973). These results indicated that the critical frequency region for the frication portion primarily consists of high-frequency components, while the critical region for the transition portion is located in a lower frequency range (i.e., where the first three formants of vowels are typically located.) One criterion for selecting a frequency range was that there would be no point on a normal-hearing subject's PA function representing a large percentage of audibility corresponding to a near-chance recognition score. The critical frequency ranges were chosen as 1500-5000 Hz for the frication and 250-3000 Hz for the transition. These frequency ranges were used consistently throughout this study.

Figure 3 displays the PA functions calculated using these frequency ranges from the normal-hearing subject's data. As shown in Figure 3, substantial recognition occurred for cases (20 and 30 dB SPL) in which 0% audibility of the frication cue (filled round symbols), and 20–50% of the transition cue (open triangular symbols), were obtained by the subjects.

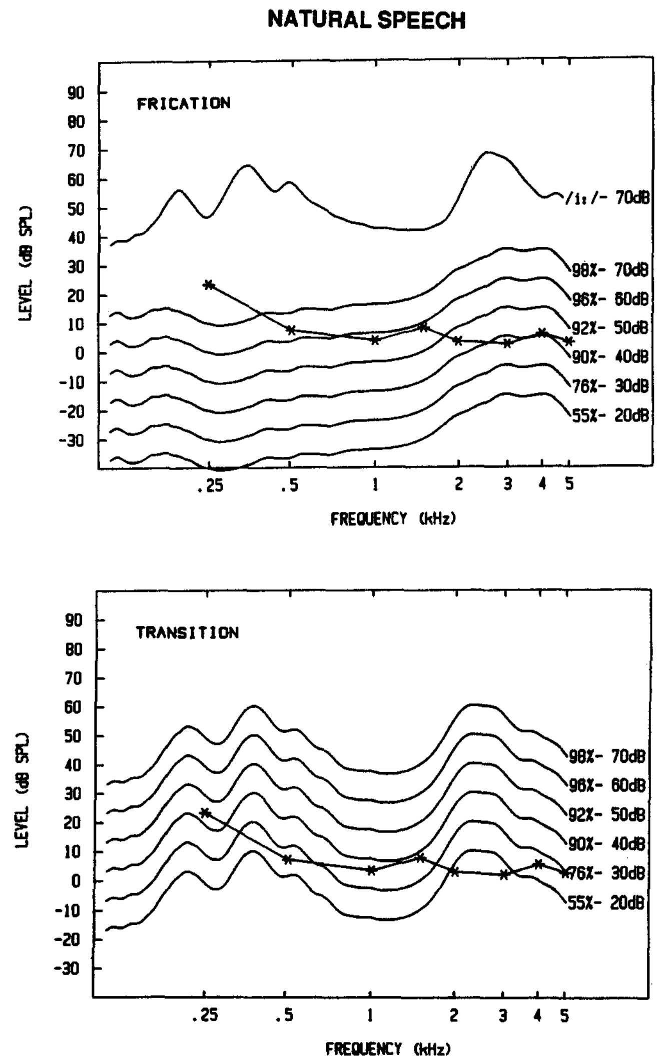

FIGURE 2. Upper panel: the mean *frication* audibility spectra of the natural stimuli at several presentation levels plotted with the mean sensitivity thresholds (*——*) for the 4 normal-hearing subjects. Percent-correct recognition and presentation level (in dB SPL) for each condition are shown to the right of the curve. Lower panel: same demonstration for the mean *transition* spectra of the natural stimuli.

Recognition Results—Condition 2

The recognition performance by the normal-hearing subjects for the synthetic stimuli employed in Condition

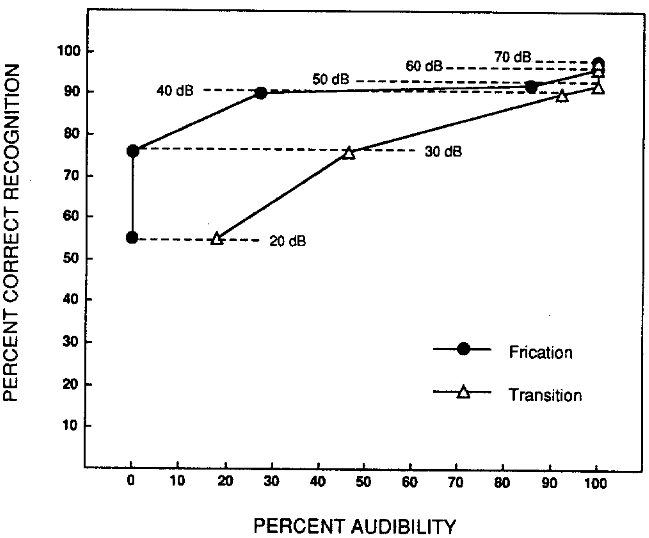

FIGURE 3. PA functions, illustrating percent-correct recognition as a function of the percent audibility of the frication and transition portions for the 4 normal-hearing subjects. The relationship between the percentage audibility and its presentation level is indicated by the dashed line.

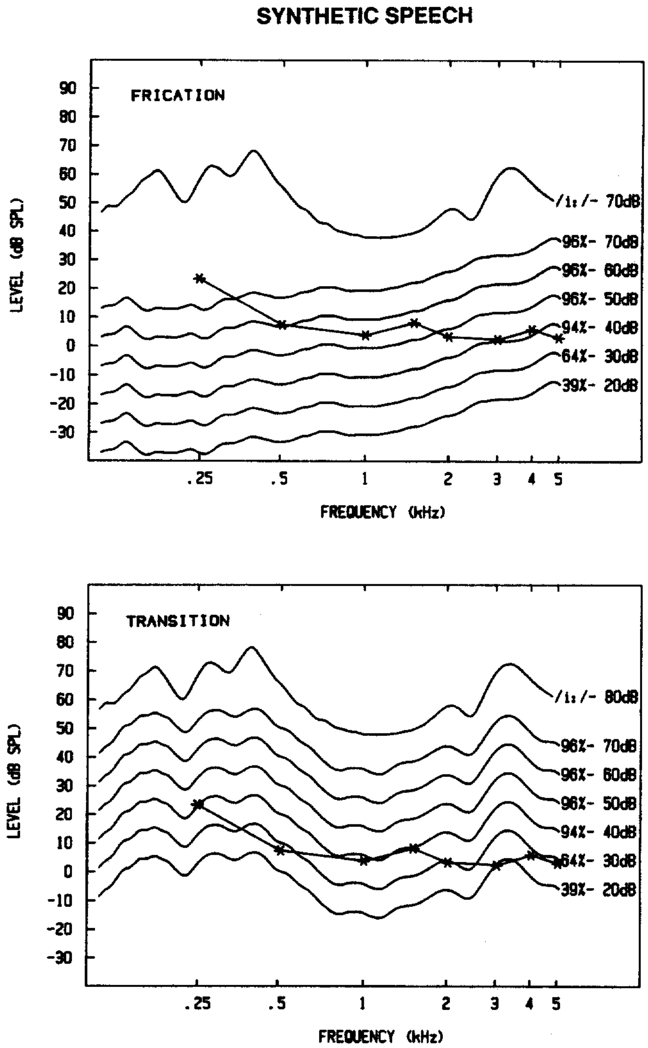

FIGURE 4. Same representation as in Figure 2 for synthetic speech in which both frication and transition cues were presented.

2 is plotted in Figure 4. It can be seen in Figure 4 that the 39%-correct recognition score obtained using the synthetic stimuli was slightly above chance performance (25% correct) when part of the audibility spectrum for transitions was just above the subjects' thresholds (at 20 dB SPL), and considerable recognition (64%) occurred as more transition cue was made audible. All of the frication spectra remained inaudible at 30 dB SPL.

The recognition scores were similar to, though slightly lower than, those found for the natural speech at the equivalent nominal presentation levels. The reason for this may be related to the removal of the additional amplitude and duration cues in the synthetic stimulus set.

Recognition Results—Condition 3

The experiment using the stimuli in Condition 3 (only the frication cue presented) was conducted to test whether the frication cue alone can serve as a sufficient cue for correct fricative recognition. The percent-correct recognition scores along with the audibility spectra are displayed in Figure 5. Recognition scores in Figure 5 clearly indicate that the frication portion alone, when audible, served as a sufficient cue for fricative perception. Specifically, subjects obtained only chance performance at 30 dB SPL when no portion of the frication was audible, but 62% correct was obtained when only a very small high-frequency portion was audible. In addition, it is also obvious from Figure 5 that the subjects could achieve almost perfect recognition (92%) when only the frequency components above 1500 Hz of the mean frication audibility spectrum were audible (at 50 dB SPL nominal presentation level). This result provides further evidence that the frequency components below 1500 Hz in the frication portion do not contribute much to the recognition (in this closed-set task) of the truncated voiceless fricatives.

In order to display the contribution of the transition cues, the performance-intensity (PI) functions obtained from the truncated fricatives, along with those from the nontruncated tokens, are shown in Figure 6. Subjects achieved only chance-level recognition performance when truncated frication stimuli were presented at the nominal level of 30 dB SPL (28% correct), whereas they obtained 64% correct recognition when the nontruncated fricatives were presented at the same level (the second leftmost filled circle). Since the truncated stimuli contained only the voiceless frication cue, whereas the nontruncated stimuli contained both frication as well as voiced transition cues, the relative importance of the transition cue to fricative recognition becomes ap-

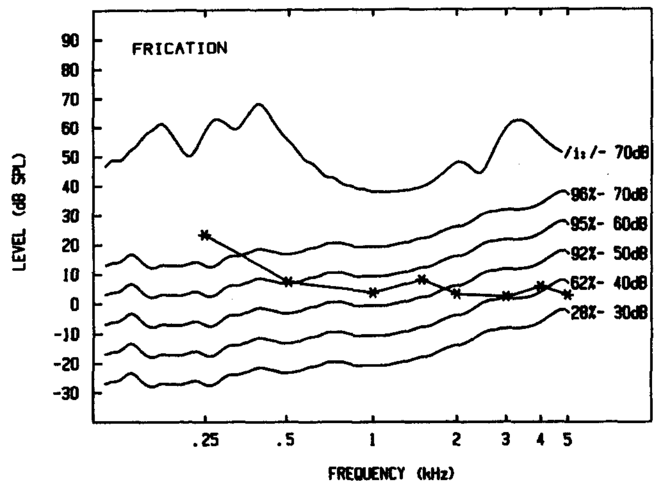

FIGURE 5. Same representation as in Figure 2 for the truncated synthetic speech in which only the frication cue was presented.

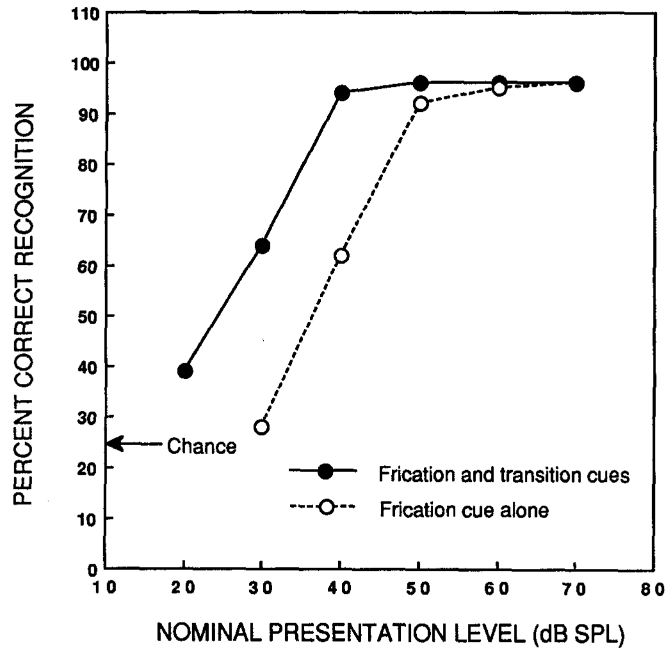

FIGURE 6. PI functions using the truncated fricative stimuli are plotted in contrast to those of the nontruncated fricative-vowel stimuli by the same normal-hearing subjects.

parent, especially when stimuli are presented at low intensity levels (20, 30, and 40 dB SPL in the current conditions).

DISCUSSION

The results from the experiment using natural speech suggested that the transition cue was used by normal-hearing subjects for recognizing the voiceless fricatives at low presentation levels where the frication cue was not usually audible. At higher presentation levels, it is possible that both frication and transition cues contribute to fricative recognition. These suggestions were further supported by the similar results obtained when using the synthetic stimuli, in which not only was there an identical vowel in all four syllables, but also the starting and ending points for frication, voicing and transition were accurately timed.

The almost perfect recognition performance at higher presentation levels using the truncated stimulus set in Condition 3 indicated that the frication cue alone can serve as a sufficient cue for correct recognition of voiceless fricatives. By comparing the results between conditions when both cues were presented and when only the frication cue was presented, the possible contribution of the transition cue, especially at low presentation levels, was demonstrated.

It seemed natural to conduct another similar experiment to further investigate whether the transition cue, by itself, could serve as a sufficient cue for fricative recognition. We performed several pilot experiments, but failed to obtain any satisfactory results. The apparent reason for the failure was that the tokens, either natural or synthetic stimuli, lost their "fricative" quality and became "stop-like" syllables when the frication portion was removed. For example, when truncated prior to the onset of voicing, the alveolar fricative transition [s] sounded like its voiced stop counterpart [d]; the labiodental [f] sounded like the labial stop [b]; the linguadental [θ] sounded something like the voiced fricative counterpart [ð] or the voiced stop [d]; the palatal [ʃ] resembled something between a palatal-velar stop [g] and glottal fricative [h]. In other words, a fricative-vowel token, when presented with only the transition cue, sounded more like the stop consonant having the closest place-of-articulation to that fricative. Due to this perceptual ambiguity, it was difficult for us to decide upon an appropriate recognition response criterion; that is, should we tell the subjects to listen for a set of fricative-vowel syllables without frication portions, or should we just tell them to recognize four stop consonants? The first approach failed because of the missing fricative quality; that is, the task became problematic for subjects who were asked to choose a fricative from a set of stop-like stimuli. The second approach was also not successful since there are only three places of articulation in normal stop consonants, and it was too confusing for subjects to distinguish between four places of stop-like articulation. Because neither approach was considered a good speech recognition task, our conclusion regarding the perceptual role of the transition cue must be limited to the conservative statement that the addition of the transition resulted in an improvement in recognition as compared to frication-alone stimuli, especially at low presentation levels.

In summary, the recognition results from both natural and synthetic stimuli by normal-hearing subjects suggest that the frication cue alone may be a sufficient cue for distinguishing the place of articulation of fricatives, whereas the transition cue, though it may not be absolutely necessary, plays an assistive role, especially at lower presentation levels.

EXPERIMENT 2

METHOD

Subjects

Three hearing-impaired adults served as subjects in this experiment. Each of the hearing-impaired subjects was diagnosed as having a cochlear hearing loss, based upon the absence of an air-bone gap and results of stapedial reflex-decay and tone-decay testing. Hearing-impaired subject AL, age 65, had a flat, severe hearing loss and had worn binaural hearing-aids for more than 5 years; HB, age 62, had a sloping bilateral severe hearing loss and had worn a hearing aid in his better ear for 10 years; KQ, age 45, had a binaural flat, moderate hearing loss and had never worn a hearing-aid. The test ear for the hearing-impaired subjects was chosen as the one with the most sensitive thresholds. Their pure-tone thresholds, displayed in Figure 7, were also determined by the same method as for the normal-hearing subjects.

Stimuli

Both types of synthetic stimuli as described in Experiment 1 were used in this experiment, but generally presented at higher intensity levels than for normal-hearing subjects. The presentation range for each hearing-impaired subject was chosen from the lowest level at which the subjects achieved a close-to-chance recognition score (25%) to the level at which the subject either achieved 100% audibility (e.g., subject KQ at 100 dB SPL) or the subject reported loudness discomfort (e.g., subjects AL and HB at 110 dB SPL). All the other procedures were identical to those described in Experiment 1.

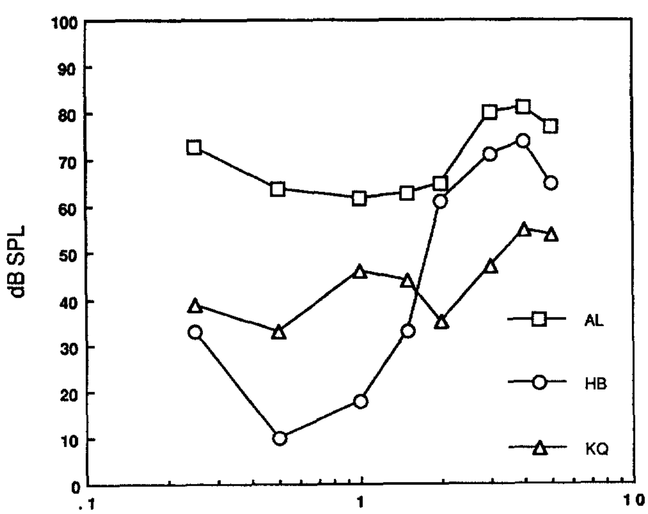

FIGURE 7. Pure-tone thresholds for the 3 hearing-impaired subjects.

RESULTS

Recognition Results—Condition 2

In order to compare recognition performance between normal-hearing and hearing-impaired subjects for equivalent degrees of audibility, results obtained in both groups were transformed to PA functions based upon the mean spectra of the four tokens. The PA functions for the 3 hearing-impaired subjects are shown in Figure 8. The frication panel displays the percent-correct recognition performance for normal-hearing and hearing-impaired subjects as a function of the audibility of the frication portion, whereas the transition panel displays recognition as a function of the audibility of the transition portion.

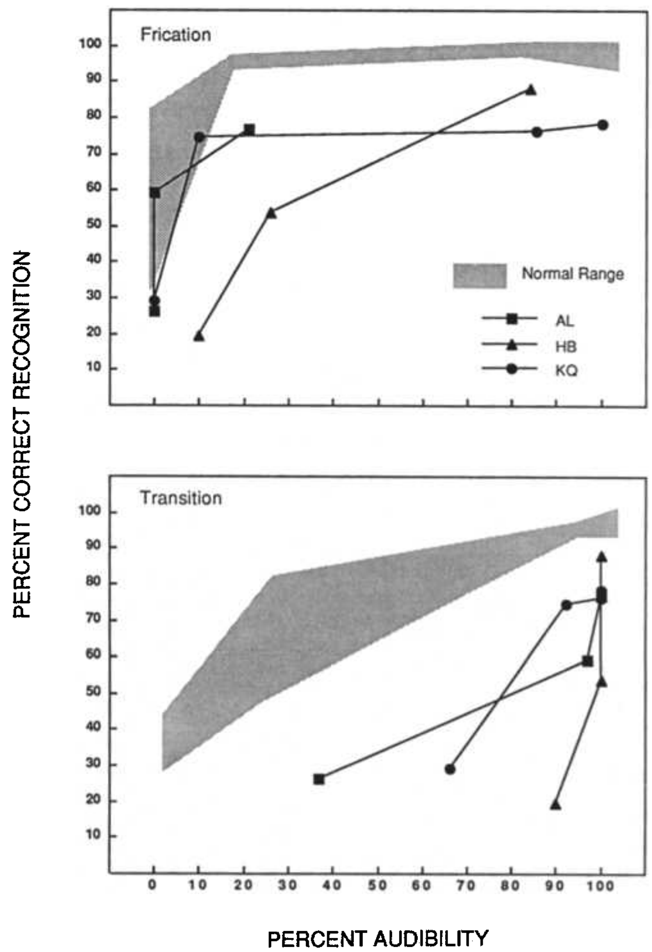

FIGURE 8. Upper panel: PA functions for the frication portion of the synthetic fricative-vowel stimuli by normal-hearing subjects are plotted in contrast to that of the hearing-impaired subjects. Lower panel: same as upper panel except for the transition portion of the same stimuli.

It can be seen in the transition panel of Figure 8 that the 3 hearing-impaired subjects were not able to use audible transition cues as efficiently as the normal-hearing subjects did. Compared to the normal-hearing subjects' results (their recognition performance range is plotted as the shaded area), the recognition scores for the hearing-impaired subjects were well below the normal range for the equivalent audibilities. For example, AL, HB, and KQ only achieved 26%, 20%, and 29% (close to chance level) correct recognition when their percentages of the transition audibility were 37%, 90%, and 66% (the leftmost filled square, circle, and triangle points in the transition panel) respectively. However, the finding that hearing-impaired subjects couldn't utilize the transition cue as efficiently as the normal subjects does not imply that they are totally lacking in this ability. One of the subjects, AL, was able to use transition cue to increase his recognition score from 26% to 59% as the transition audibility increased from 37% to 97%, whereas the frication was totally inaudible (see the first two leftmost filled square points in both panels of Figure 8).

Since hearing-impaired subjects were poor users of the transition cue, their recognition of the fricative-vowel syllables in this experiment appeared to come primarily from the contribution of the frication cue. If this were the case, then equivalent frication audibility would result in equivalent recognition performance. However, testing this hypothesis using the data in the frication panel of Figure 8 is impossible because the recognition results, though displayed as a function of the audibility of the frication cue, included the contribution of the transition cue, which was more efficiently used by normal-hearing subjects. One additional factor was that at the highest presentation level, hearing-impaired subject AL reached his loudness discomfort level, while his audibility for the frication portion was still very low (about 30%, the rightmost filled square in the frication panel).

Recognition Results—Condition 3

The hypothesis that fricative recognition by hearing-impaired subjects is due primarily to the contribution of the frication cue was directly tested in this experiment. The nominal presentation levels used for hearing-impaired subjects were 100, 110, 120 dB SPL for AL; 90, 100, 110 dB SPL for HB; and 70, 80, 90, 100 dB SPL for KQ. It is important to remember that the 120 dB SPL nominal presentation level for AL in this experiment meant that the actual frication portions of CV tokens were presented at approximately 90 dB SPL, because all nominal presentation levels were referenced to the absent vowel intensity. These results are converted to PA functions (in terms of the mean frication audibility spectra) and displayed in Figure 9, in which the normal-hearing subjects' recognition performance range under the same condition was included for comparison. Figure 9 shows that hearing-impaired subjects achieved essentially normal recognition performance when given equivalent degrees of audibility of the frication cue. It can also be seen

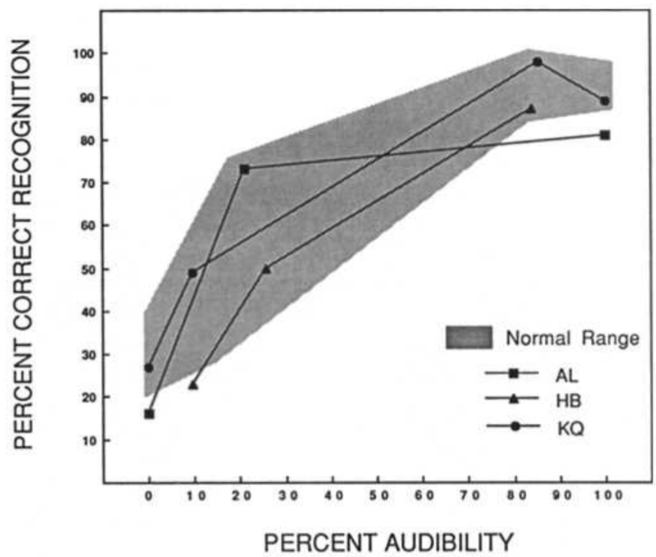

FIGURE 9. PA functions using the truncated fricative stimuli obtained by both normal-hearing and hearing-impaired subjects.

in Figure 9 that hearing-impaired subject KQ's recognition score was improved by 10%–20% at high presentation level when compared with the experiment in which both frication and transition cues were presented (Figure 8). We do not know the reason for this.

The result that under these controlled conditions the hearing-impaired subjects performed as well as normal-hearing subjects, is also considered an important finding of the current study. Since hearing-impaired subjects essentially achieved perfect recognition at a high percentage of audibility, this conclusion is unchanged even if the hearing-impaired subjects had broader than normal summating filter bandwidths (Higgins & Turner, 1988), which would result in greater audibility than calculated here for the frication portions of the consonants. The importance of this finding resides in the fact that hearing-impaired subjects can still utilize their residual auditory ability to discriminate the gross spectral difference between fricatives. This seems reasonable for the hearing-impaired subjects, if we speculate that only minimal frequency selectivity is required for the discrimination of the gross spectral shapes of the steady-state frication portion of fricatives.

DISCUSSION

Results of both experiments, with normal and hearing-impaired subjects, suggest that the steady-state spectral cues in the frication portion may be sufficient for fricative recognition by both groups of subjects. In contrast, the dynamic spectral cues present in the formant transition, although not absolutely necessary, can assist normal-hearing subjects to distinguish place information, especially at low presentation levels. Hearing-impaired subjects appear to utilize the transition cue in some cases;

however, they do not seem to use it as efficiently as normal-hearing subjects.

These findings imply that a primary factor in hearing-impaired subjects' difficulty in distinguishing fricatives is not only the elevation of their hearing thresholds and the low intensities of the frication cues (relative to vowel), but also their poorer-than-normal discrimination abilities for the dynamic spectral cues occurring in the formant transition between the fricative and the subsequent vowel. Consider a typical case (Figure 10) of moderate to severe hearing impairment (Subject HB), where the elevation of thresholds has reduced the available dynamic range (dark area) to 30 dB at high frequencies. When the vowel portion (V) has been amplified to reach the loudness discomfort level, the frication portions (F) would still be inaudible to the hearing-impaired subjects, because they are typically about 30 dB lower than the vowel in intensity. On the other hand, although the transition portions (T), which are about 5–10 dB less intense than the vowel, would be audible to the hearing-impaired subject, he is unable to make full use of them.

To conclude, the difficulties that hearing-impaired subjects have in recognizing voiceless fricative consonants may be due to the fact that the frication cue, which they can utilize as well as normal subjects, is often inaudible to them, while at the same time they cannot make efficient use of the transition cue, which is usually audible. In other words, both the lack of audibility and the loss of discriminability appear to be primary factors responsible for hearing-impaired subjects' difficulties in fricative recognition. However, the lack of audibility may be an easier issue to solve in the sense that the frication cue might be well-perceived by hearing-impaired listeners if it can be made audible by appropriate amplification, as opposed to the transition cues, which cannot be compensated for by similar amplification.

If the present explanations for the hearing-impaired subjects' difficulty in perceiving voiceless fricatives are correct, we would expect that any effort to increase the audibility of the frication portion will result in the improvement of fricative intelligibility. Although more investigations are needed to further validate this conclusion, hearing-aid designers may have to consider the lack of audibility for weak intensity consonants. In other words, a short-term rms level for consonants may have value when considering hearing-aid amplification characteristics.

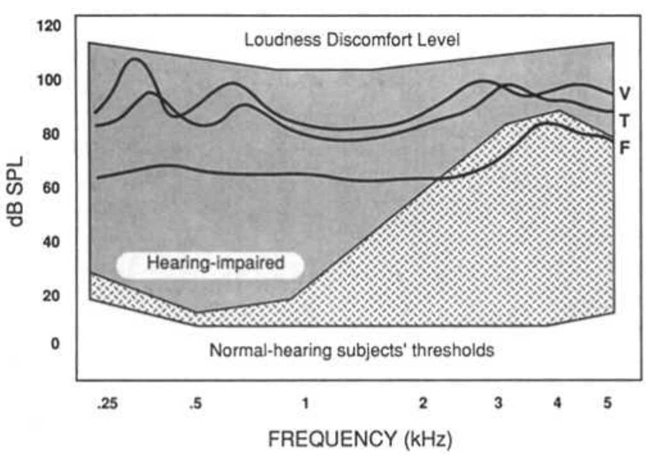

FIGURE 10. Schematic illustration of the difficulties that hearing-impaired subjects have in fricative recognition (see text).

ACKNOWLEDGMENTS

This research was supported in part by a grant from the Syracuse University Senate Research Grant Awards, the Deafness Research Foundation, and NIH grant NS24175. The authors thank Edward Conture for a detailed and critical reading of an earlier version of this manuscript. We also thank the associate editor, Arlene Carney, and the other two reviewers for their helpful suggestions in organizing the final version of the manuscript.

REFERENCES

BEHRENS, S., & 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. *Journal of the Acoustical Society of America, 84,* 861–867.

BILGER, R., & WANG, M. (1976). Consonant confusions in patients with sensorineural hearing loss. *Journal of Speech and Hearing Research, 19,* 718–748.

FANT, G. (1960). *Acoustic theory of speech production.* Mouton, 's-Gravenhage: The Netherlands.

FLETCHER, H. L. (1953). *Speech and hearing in communication.* Van Nostrand: New York.

GÄSSLER, G. (1954). Uber die Horschwelle fur Schallereignisse mit Verscheiden breitem Frequenzspktrum. *Acustica, 4,* 408–414.

GREEN, D. M. (1958). Detection of multiple component signals in noise. *Journal of the Acoustical Society of America, 30,* 904–911.

GUERLEKIAN, J. A. (1981). Recognition of the Spanish fricatives /s/ and /f/. *Journal of the Acoustical Society of America, 70,* 1624–1627.

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

HEINZ, J. M., & STEVENS, K. N. (1961). On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America, 33,* 589–596.

HIGGINS, M. B., & TURNER, C. W. (1988, November). *Summating bandwidths at threshold of normal and hearing-impaired listeners.* Paper presented at the Convention of American Speech-Language/Hearing Association, Boston, MA.

HUGHES, G. W., & HALLE, M. (1956). Spectral properties of fricative consonants. *Journal of the Acoustical Society of America, 28,* 303–310.

JONGMAN, A. (1989). Duration of frication noise required for identification of English fricatives. *Journal of the Acoustical Society America, 85,* 1718–1725.

KLATT, D. H. (1980). Software for cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America, 67,* 971–995.

LEVITT, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49,* 467–477.

MCCASLAND, G. P. (1979a). Noise intensity and spectrum cues for spoken fricatives. *Journal of the Acoustical Society of America, Suppl. 1, 65,* S78–S79.

MCCASLAND, G. P. (1979b). Noise intensity cues of spoken fricatives. *Journal of the Acoustical Society of America, Suppl. 1, 66,* S88.

MINIFIE, F. D., HIXON, T. J., & WILLIAMS, F. (1973). *Normal aspects of speech, hearing and language*. Englewood Cliffs, NJ: Prentice-Hall.

OWENS, E. (1978). Consonant errors and remediation in sensorineural hearing loss. *Journal of Speech and Hearing Disorders, 43*, 331–347.

OWENS, E., BENEDICT, M., & SCHUBERT, E. D. (1972). Consonant phonemic errors associated with pure tone configurations and certain kinds of hearing impairment. *Journal of Speech and Hearing Research, 15*, 308–322.

SPIEGEL, M. F. (1979). The range of spectral integration. *Journal of the Acoustical Society of America, 66*, 1356–1363.

STEVENS, K. N. (1971). Airflow and turbulence noise for fricative and stop consonants: Static considerations. *Journal of the Acoustical Society of America, 50*, 1180–1192.

STEVENS, K. N. (1985). Evidence for the role of acoustic boundaries in the perception of speech sounds. In V. A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged*. (pp. 243–255) Academic: New York.

STREVENS, P. (1960). Spectra of fricative noise in human speech. *Language and Speech, 3*, 32–49.

TURNER, C. W., & ROBB, M. P. (1987). Audibility and recognition of stop consonants in normal and hearing-impaired subjects. *Journal of the Acoustical Society of America, 81*, 1566–1573.

VAN TASELL, D. J. (1981). Auditory perception of speech. In J. M. Davis (Ed.), *Rehabilitative audiology for children and adults*. (pp. 13–59) New York: John Wiley & Sons, Inc.

Received April 7, 1989
Accepted January 29, 1990

Requests for reprints should be sent to Fan-Gang Zeng, Communication Sciences and Disorders, Syracuse University, 805 South Crouse Avenue, Syracuse, NY 13244–2280.