Research Article

Effects of Signal Level and Spectral Contrast on Vowel Formant Discrimination for Normal-Hearing and Hearing-Impaired Listeners

Ashley Woodall^a and Chang Liu^a

Purpose: The aim of this study was to determine whether increasing the overall speech level or the individual spectral contrasts of vowel sounds can improve vowel formant discrimination for listeners both with and without normal hearing.

Method: Thresholds of vowel formant discrimination were examined for the F2 frequencies of 3 American English vowels for listeners with and without normal hearing. Spectral contrasts of the F2 were enhanced by 3, 6, and 9 dB. Vowel stimuli were presented at 70 and 90 dB SPL.

Results: The thresholds of listeners with hearing impairment were reduced significantly after spectral

enhancement was implemented, especially at 90 dB SPL, whereas normal-hearing listeners did not benefit from spectral enhancement.

Conclusion: These results indicate that a combination of spectral enhancement of F2 and high speech level is most beneficial to improve vowel formant discrimination for listeners with hearing impairment.

Key Words: speech perception, hearing loss, speech enhancement

Listeners with hearing impairment (HI) have greater difficulty in speech perception than normal-hearing (NH) listeners (Henry, Turner, & Behrens, 2005; Leek, Dorman, & Summerfield, 1987). Because poor communication performance for listeners with HI is often caused by reduced audibility of speech signals, a common method of compensating for hearing loss is to increase the intensity of speech sounds beyond audibility levels for listeners with HI (Humes, Dirks, Bell, & Kincaid, 1987). However, reduced frequency selectivity is often associated with sensorineural hearing loss and contributes to difficulty in speech perception for listeners with HI (Needleman & Crandell, 1997). Therefore, another compensatory method is to selectively enhance the spectral and temporal features that are important for speech perception (i.e., increase the spectral contrast of speech stimuli; Bunnell, 1990; Lyzenga, Festen, & Houtgast, 2002; Stone & Moore, 1992). The primary purpose of this study was to investigate whether increasing the speech level or enhancing the spectral contrasts of vowel sounds can improve vowel formant discrimination for listeners with HI.

A number of studies have suggested that vowel sounds are perceptually categorized by their characteristic spectral properties, such as the resonant peaks in the vowel spectrum (Hillenbrand, Getty, Clark, & Weeler, 1995). These peaks, called formants—especially the first two formants, F1 and F2—are primary acoustic cues for vowel identification and categorization. Small changes to vowel formant frequency may result in a perceptual shift from one vowel category to another. Therefore, it is important to understand how sensitive the human auditory system is to differences in formant frequency. The threshold for vowel formant discrimination is defined as the smallest detectable difference of a single formant frequency. Many factors have been shown to significantly affect formant discrimination for NH listeners, including vowel category, level of linguistic context, and speech level (Liu, 2008; Liu & Kewley-Port, 2004).

To date, studies of vowel formant discrimination for listeners with HI are limited, providing equivocal results. Coughlin, Kewley-Port, and Humes (1998) and Richie, Kewley-Port, and Coughlin (2003) measured formant discrimination for four American English vowels presented at conversational levels (60–70 dB SPL) and relatively high sound levels (95 dB SPL) for NH and HI listeners. Results suggested that, for the listeners with HI, formant discrimination improved significantly from low to high levels, especially for F2 frequency, in which the listeners with HI had hearing loss. These two studies argued that this improvement was due to the better audibility at high levels.

^aThe University of Texas at Austin

Correspondence to Chang Liu: changliu@utexas.edu

Editor: Larry Humes

Received July 27, 2012

Revision received October 30, 2012

Accepted October 31, 2012

DOI: 10.1044/1059-0889(2012)12-0044)

However, the listeners with HI still had markedly higher thresholds than the NH listeners for F2 at 95 dB SPL, indicating that increased audibility alone was not sufficient for the listeners with HI to achieve NH performance.

Conversely, Liu and Kewley-Port (2007) reported that formant discrimination of F2 at a well-audible level (95 dB SPL) for listeners with HI was worse than at 70 dB SPL in three phonetic contexts: isolated vowel, syllable, and sentence. They argued that this reverse-level effect on formant discrimination was due to the reduced frequency selectivity and upward spread of masking on F2 produced by F1. Additionally, Liu and Kewley-Port concluded that in the earlier studies (Coughlin et al., 1998; Richie et al., 2003), F2 was not audible at low levels and became well audible at high levels, whereas in their study, F2 was audible at both the low and high levels. Together, these studies with listeners with HI suggest that formant discrimination was significantly affected not only by audibility, but also by other factors associated with hearing loss such as reduced frequency selectivity.

In addition to the studies that measured vowel formant discrimination of listeners with HI, other studies examined just-noticeable differences in F2 transition of listeners with HI, with a main focus on the effect of upward spread masking of F1 on F2 discrimination (Danaher, Osberger, & Pickett, 1973; Danaher & Pickett, 1975; Van Tassell, 1980). Pickett and colleagues (Danaher et al., 1973; Danaher & Pickett, 1975) found that discrimination of F2 transitions in synthetic vowels was improved for listeners with moderate-to-severe flat or sloping hearing losses as F1 was removed from the stimuli, indicating that the upward spread of masking from F1 interfered with the F2 discrimination of the listeners with HI. In a later study, Van Tassell (1980) reported that as a group, listeners with HI performed similarly in F2 transition discrimination for stimuli with and without F1 presented, although the listeners with HI showed great individual variability. These studies indicate that the presence of low-frequency components (e.g., F1) generates masking on F2 perception. Thus, spectral enhancement of F2 in the present study may increase the F2 representation to overcome the upward spread of masking by F1.

Based on these findings of F2 discrimination (Coughlin et al., 1998; Danaher et al., 1973; Danaher & Pickett, 1975; Liu & Kewley-Port, 2007; Richie et al., 2003; Van Tassell, 1980), we proposed that the enhancement of specific acoustic features of vowels may improve discrimination for listeners with HI rather than only increasing overall vowel intensity. One such solution is to enhance spectral contrasts. Vowel formants are represented with much lower spectral contrast in the dysfunctional auditory system of humans (Liu & Kewley-Port, 2007) and animals (Miller, Calhoun, & Young, 1999b) as compared to the normal auditory system. Several studies have suggested that vowel identification improves as the spectral contrast of vowel formants are increased for both NH and listeners with HI, indicating the importance of spectral contrast for vowel perception (Leek et al., 1987; Liu & Eddins, 2008). Increasing the spectral contrasts of F2 and F3 improved the spectral representation of vowels and partially restored sensitivity to the F2 frequency in the

responses of the HI auditory nerve (Miller, Calhoun, & Young, 1999a). Thus, spectral enhancement of vowel formants may be able to compensate for the spectral smearing that is caused by sensorineural hearing loss. Spectral enhancement may also improve vowel formant discrimination for listeners with HI.

In the present study, we increased the spectral contrasts of F2 for individual vowels. This is in contrast to the high-pass amplification of the F2 and F3 frequency ranges that are commonly used in hearing aid technology by both manufacturers and traditional prescriptive fitting methods (Mueller, 2005). When high-pass amplification is applied, the spectral valley between F1 and F2 could be increased, obscuring the spectral representation and worsening listener performance (Miller et al., 1999a). Thus, in this study, we enhanced the spectral contrasts of F2 by increasing the amplitude of F2 but not changing the spectral valleys near F2. If the spectral enhancement of selective formant peaks is found to improve speech perception for listeners with HI, this strategy may have potential implications in signal processing algorithms for amplification devices.

Altogether, the main purpose of this study was to investigate whether acoustic manipulation, via increase of the overall speech level or spectral contrast, is able to improve vowel formant discrimination for NH and HI listeners. We manipulated three factors: speech level (70 and 90 dB SPL), spectral enhancement of F2 (original and enhancement by 3, 6, and 9 dB), and F2 frequency of three English vowels (/Λ, æ, i/).

Method

Stimuli

Three American English isolated vowels, /Λ, æ, i/, were used as stimuli. The three original vowels were recorded in an /hVd/ context from a young American English female talker and served as the base stimuli for spectral enhancement. Isolated vowels were obtained by truncating the initial /h/ and the final /d/, including the onset and offset formant transition, such that only the central vowel nucleus remained. The central 200-ms portions of the /hVd/ syllable that included only vowels in isolation were selected as the base stimuli for spectral enhancement and formant frequency shifts. Because previous studies (Coughlin et al., 2008; Liu & Kewley-Port, 2007; Richie et al., 2003) reported that listeners with HI primarily had difficulty discriminating differences of F2 frequency, the present study focused on formant discrimination of F2. The average F2 frequency of the three vowels over the entire vowel duration was 1442 Hz for /Λ/, 2078 Hz for /æ/, and 2563 Hz for /i/.

There were four conditions of formant enhancement: unenhanced (original), and 3-, 6-, and 9-dB enhancement of F2. The left panel of Figure 1 illustrates the spectra of the original /æ/ vowel (solid line) and the /æ/ vowel with 9-dB F2 enhancement (dashed line). The spectral enhancement of F2 was manipulated as follows: first, a three-dimensional (3-D) spectrogram (Amplitude × Time × Frequency) of the original vowel was obtained by analysis using STRAIGHT

Figure 1. Left panel: spectra of the unenhanced vowel /æ/ (solid line) and the vowel /æ/ with 9-dB F2 enhancement (dashed line) in STRAIGHT with a sampling frequency of 12207 Hz. Right panel: spectra of the unenhanced vowel /æ/ (solid line) and vowel /æ/ with 10% F2 shift (dashed line) in STRAIGHT with a sampling frequency of 12207 Hz.

(Kawahara, Masuda-Kastuse, & Cheveigne, 1999); second, at each time frame of the spectrogram, the valleys below F1, between F1 and F2, and between F2 and F3 were defined as V1, V2, and V3, respectively. For the enhanced F2 conditions, only amplitudes at the frequencies between the valleys next to F2 (e.g., V2 and V3) were enhanced. The level of the F2 peak was amplified by 3, 6, and 9 dB, and the levels of the components between V2 and F2 and between F2 and V3 were linearly interpolated in a log scale. For example, the amplitude at the midpoint of V2 and F2 was amplified for 4.5 dB for the 9-dB enhancement conditions. Thus, there were four vowels with and without F2 enhancement serving as standard vowels for formant frequency shifts for each vowel category.

Stimuli were generated with different amounts of formant-frequency shift based on each standard vowel with the corresponding spectral enhancement. The procedure for shifting F2 frequency is briefly described below (for more details, see Liu & Kewley-Port, 2004). For each time frame of the 3-D spectrogram (Kawahara et al., 1999), as shown in the right panel of Figure 1, the F2 peak was shifted by flattening the frequency range corresponding to the formant shift at the spectral valley lower than the F2 frequency (V2) and replacing the original amplitude values with the shifted peak at the high-frequency valley (V3). In other words, the F2 shift was manipulated between the two spectral valleys, V2 and V3. This was also the same frequency range of the F2 enhancement. The F2-shifted vowels had the same amplitude of formant peaks (e.g., F1, F2, and F3) as the standard vowel. The F2 values were shifted systematically from 0.5% to 20% of F2 frequency, with a step size of 0.5%, resulting in 40 formant-shifted vowels. This F2-shifted spectrogram was loaded into STRAIGHT and was used for vowel resynthesis. Thus, for each vowel, there were four sets of 40 F2-shifted vowel stimuli, with each set corresponding to one F2 enhancement condition. Additionally, a vowel /ε/ was recorded from a middle-aged female speaker, different from the female speaker who recorded the three test vowels. The vowel /ε/ served as the standard vowel for a 15-min training

session before the test sessions. Like the test vowels, the F2 of the vowel /ε/ was shifted from 0.5% to 20% of F2 frequency, with a 0.5% step size with the same formant shift method described above.

The stimuli with and without spectral enhancement were presented with 10-ms rise-fall ramps at two levels: 70 and 90 dB SPL. Sound pressure levels of vowel sounds were measured at the output of the ER-2 insert earphones via an NBS-9A 2-cc coupler that was connected to the microphone of a Larson-Davis (Model 2800) sound-level meter set to the linear weighting scale.

Study Participants

Six HI listeners with moderate sensorineural hearing loss at high frequencies and six NH listeners participated in this study. Participants were 18–55 years old; NH and HI listeners had mean ages of 23 and 37, respectively. All were native speakers of American English with normal middle-ear function. As shown in Figure 2, the audiometric thresholds of the listeners with HI were no more than 65 dB HL within the range of 250–4000 Hz except participant S1 at 4000 Hz. The NH listeners had pure-tone thresholds ≤15 dB HL at octave intervals between 250 and 8000 Hz (American National Standards Institute, 2010). All procedures were approved by the Institutional Review Board of the University of Texas at Austin. All listeners consented to participate in the study.

Procedure

Speech stimuli, sampled at 12207 Hz, were presented to the right ears of the listeners, who were seated in a sound-treated IAC (Industrial Acoustics Company) booth, via calibrated ER-2 insert earphones. Stimulus presentation was controlled by a series of TDT (Tucker-Davis Technologies) modules, including an enhanced real-time processor (RP2.1), a programmable attenuator (PA5), and a headphone buffer (HB7). Thresholds were measured using a three-interval,

Figure 2. Audiograms of the six listeners with hearing impairment (HI) (right ear), from 250 to 4000 Hz, showing the configuration of hearing loss over the frequency range.

forced-choice procedure with a two-down, one-up tracking algorithm, estimating 71% correct responses (Levitt, 1971). There were 24 experimental conditions (3 formant frequencies $\times$ 4 spectral enhancements $\times$ 2 speech levels), the order of which was randomized.

For each trial, the standard vowel was presented in the first interval, followed by a standard vowel and a formant-shifted vowel randomly ordered in the second and third test intervals. The listener's task was to indicate which of the two test intervals contained the vowel that was different from the standard vowel via a button press on an LCD monitor with an interface that simulated a handheld button box. The interstimulus interval (ISI) was 400 ms. A light located above each interval button illuminated simultaneously with each presentation of the vowel stimulus. For each block, the F2 shift was started at 10% of the F2 frequency and was adjusted in a 1% step size for the first three reversals and then 0.5% thereafter. The threshold for each block of 60 trials was based on the average F2 shift at the last even number of reversals in the adaptive track, without counting the first three.

Across all conditions and listeners, a threshold of a given block was computed as a mean of 10 reversals. The threshold for each condition was taken as the average for two blocks. An additional block was conducted if thresholds for the two blocks differed by $>1\%$ of the F2 frequency (e.g., two steps of formant shifts). Across all of the listeners, this occurred an average of eight times out of a total of 24 experimental conditions. Listeners completed a 15-min training session using the vowel /e/ before the test session to gain familiarization with the procedure. The training and test sessions lasted ~ 5 hr, with each session lasting 90–120 min. The experimental design and process was manipulated using SykoFizX software application (TDT, Inc.).

Results

Overall Performance

Thresholds are expressed as Weber fractions ($\Delta F/F$). Average thresholds across NH (left) and HI (right) listeners are plotted in Figure 3 as a function of F2 frequency for the four enhancement conditions and two speech levels. A four-factor (within-subjects factors: spectral enhancement, formant frequency, and speech level, and between-subjects factor: listener group) analysis of variance (ANOVA) of the thresholds was conducted. Thresholds were significantly affected by listener group, $F(1, 10) = 8.239$, $p < 0.05$; formant frequency, $F(2, 20) = 7.722$, $p < 0.05$; and spectral enhancement, $F(3, 30) = 20.299$, $p < 0.05$, but not by speech level, $F(1, 10) = 0.158$, $p > 0.05$. Of the six two-factor interactions, four were significant (Group $\times$ Formant Frequency, Group $\times$ Spectral Enhancement, Formant Frequency $\times$ Spectral Enhancement, and Speech Level $\times$ Spectral Enhancement; all $p < 0.05$). Of the four three-factor interactions, only the Group $\times$ Speech Level $\times$ Spectral Enhancement interaction was significant ($p < 0.05$).

Because the two listener groups differed significantly in overall performance (see Figure 3), and the interaction effect of group and spectral enhancement was significant, we conducted separate analyses for the NH and HI listeners. The effects of spectral enhancement and speech level are reported in the following two subsections for NH and HI listeners, respectively.

Effects of Spectral Enhancement and Speech Level for NH Listeners

Overall, as shown in Figure 3, neither spectral enhancement nor speech level produced an improvement for the NH listeners. A three-factor (spectral enhancement, formant frequency, speech level) repeated measures ANOVA with threshold ($\Delta F/F$) as the dependent variable showed that NH listeners were not affected by speech level, $F(1, 5) = 0.098$, $p > 0.05$; spectral enhancement, $F(3, 15) = 1.482$, $p > 0.05$; or formant frequency, $F(2, 10) = 0.924$, $p > 0.05$. None of the two-way interactions for NH listeners was significant (all $p > 0.05$), but the three-way interaction of spectral enhancement, formant frequency, and speech level was significant, $F(6, 30) = 2.506$, $p < 0.05$.

Effects of Spectral Enhancement and Speech Level for Listeners With HI

A similar ANOVA for the listeners with HI showed significant effects of spectral enhancement, $F(3, 15) = 20.206$, $p < 0.05$, and formant frequency, $F(2, 10) = 6.370$, $p < 0.05$, but no significant effect of speech level, $F(1, 5) = 0.116$, $p > 0.05$. Of the two-factor and three-factor interactions, the interaction of speech level and spectral enhancement was significant, $F(3, 15) = 6.941$, $p < 0.05$, as was the interaction of formant frequency and spectral enhancement, $F(6, 30) = 2.652$, $p < 0.05$. Post hoc Tukey tests suggested that for the listeners with HI, thresholds were improved by spectral enhancement of 3, 6, and 9 dB (all $p < 0.05$). Because the

Figure 3. Average thresholds of formant discrimination in Weber fractions for normal-hearing (NH) (left) and HI (right) listeners as a function of formant frequency (F2 of /Λ/, /æ/ and /i/ from left to right) for two speech levels: 70 dB SPL (upper) and 90 dB SPL (lower), with four F2 enhancement conditions: original, 3, 6, and 9 dB.

interaction effects of spectral enhancement and speech level or formant frequency were significant, we examined the simple main effect of spectral enhancement for each speech level and for each formant frequency, described in the next two paragraphs.

In order to measure the simple main effect of spectral enhancement under each speech level for the listeners with HI, two-way (spectral enhancement and formant frequency) repeated measures ANOVAs were completed for 70 and 90 dB SPL, respectively. At 70 dB SPL, there was a significant effect of spectral enhancement, $F(3, 15) = 6.557$, $p < 0.05$, and formant frequency, $F(2, 10) = 5.565$, $p < 0.05$, on formant thresholds; there was no significant interaction effect of spectral enhancement and formant frequency, $F(6, 30) = 1.397$, $p > 0.05$. The average thresholds improved by 17% for 3-dB F2 enhancement, 24% for 6-dB F2 enhancement, and 34% for 9-dB F2 enhancement compared to the thresholds of original vowels. Post hoc Tukey tests indicated that thresholds for 6- and 9-dB enhancement were significantly better than thresholds of original vowels (both $p < 0.05$). For 90 dB SPL, spectral enhancement, $F(3, 15) = 17.015$, $p < 0.05$, and formant frequency, $F(2, 10) = 4.421$, $p < 0.05$, showed a significant effect on formant discrimination; the interaction of spectral enhancement and formant frequency did not, $F(6, 30) = 1.544$, $p > 0.05$. Thresholds improved by 46% for 3-dB F2 enhancement, 60% for 6-dB F2 enhancement, and 71% for 9-dB F2 enhancement relative to the thresholds of original vowels. Post hoc Tukey tests reported that performance was significantly improved for all three spectral enhancement conditions compared to performance of the original condition (all $p < 0.05$).

Similarly, in order to measure the main effect of spectral enhancement under each formant frequency, we conducted two-factor (spectral enhancement and speech level) repeated measures ANOVAs for the three formant frequencies. At all three formant frequencies, thresholds were significantly affected by spectral enhancement (all $p < 0.05$), but not by speech level (all $p > 0.05$). The interaction effect of speech level and spectral enhancement was significant for the F2 of /æ/ and /i/ (both $p < 0.05$), but not for the F2 of /Λ/. Post hoc Tukey tests suggested that thresholds of any of the three spectrally enhanced conditions were significantly lower than thresholds for the original conditions of the three formant frequencies (all $p < 0.05$). As shown in Table 1, individual variability of formant discrimination performance due to spectral enhancement was observed for the listeners with HI, possibly due to the different configurations and severities of hearing loss across them.

It should also be noted that for the listeners with HI, an increase of speech level did not improve formant discrimination of original vowels with no F2 enhancement (see Figure 3). For instance, for original vowels with no spectral enhancement, thresholds degraded by 34% at 90 dB SPL compared to thresholds at 70 dB SPL, whereas for 3-dB, 6-dB, and 9-dB F2 enhancement, an increase of speech level from 70 to 90 dB SPL improved formant discrimination by 14%, 31%, and 42%, respectively. These results imply that benefits from 70 to 90 dB SPL, if any, were conditional upon F2 amplitude.

As described in the statistical results above, vowel formant discrimination was not affected by speech level but was significantly affected by the interaction of spectral enhancement and speech level. To reveal the simple main

Table 1. Average improvement of formant discrimination thresholds (in percentage) over the three vowels and two speech levels for the three F2 enhancement conditions. The last row shows the average improvement over the three F2 enhancement conditions.

F2 condition	Participant					
	S1	S2	S3	S4	S5	S6
3 dB	32%	58%	12%	47%	3%	30%
6 dB	44%	71%	32%	69%	15%	20%
9 dB	54%	79%	43%	71%	27%	38%
Average	43%	69%	29%	62%	15%	29%

effect of speech level under each spectral enhancement condition, we completed two-factor (speech level and formant frequency) repeated measures ANOVAs for the vowels with original spectrum, and 3-dB, 6-dB, and 9-dB F2 enhancement, respectively. Results showed no significant effect of speech level at any of the F2 enhancement conditions: original, $F(1, 5) = 2.005$, $p > 0.05$; 3 dB, $F(1, 5) = 0.321$, $p > 0.05$; 6 dB, $F(1, 5) = 0.648$, $p > 0.05$; and 9 dB, $F(1, 5) = 2.158$, $p > 0.05$. In addition, no significant interaction effects of speech level and formant frequency were found for any of the F2 enhancement conditions: original, $F(2, 10) = 0.853$, $p > 0.05$; 3 dB, $F(2, 10) = 2.440$, $p > 0.05$; 6 dB, $F(2, 10) = 0.428$, $p > 0.05$, and 9 dB, $F(2, 10) = 0.575$, $p > 0.05$.

NH Listeners Versus Listeners With HI

As reported earlier, the HI listeners' overall performance was significantly worse than that of the NH listeners. In addition, we conducted three-factor (within-subjects factors: formant frequency and spectral enhancement; between-subject factors: listener group) ANOVAs at 70 and 90 dB SPL, respectively, indicating significantly worse performance for the listeners with HI than their NH peers (both $p < 0.05$). To determine if the increase of speech level and spectral enhancement of F2 could improve HI listeners' thresholds to NH listeners' level, we conducted a two-factor (formant frequency and listener group) ANOVA: the 9-dB F2 enhancement condition at 90 dB SPL that provided the greatest benefits for listeners with HI (e.g., the dashed line in the lower right panel of Figure 3) versus the original, unenhanced vowels at 70 dB SPL for NH listeners (e.g., the solid line in the upper left panel of Figure 3). The NH listeners showed significantly lower thresholds than the listeners with HI, $F(1, 10) = 5.054$, $p < 0.05$, suggesting that increasing speech level and spectral contrast were not able to help the listeners with HI reach the same level of performance as the NH listeners.

Discussion

The purpose of this study was to examine whether enhancing a crucial spectral prominence of vowel stimuli would improve NH and HI listeners' abilities to discriminate differences in F2 frequency. As shown in the left panels of Figure 3, the NH listeners did not show any significant

improvement from increasing speech levels or spectral enhancements of F2. This is not surprising, because NH listeners do not have the reduced amplitude and secondary distortion associated with hearing loss. NH listeners only need 3- to 4-dB peak-to-valley contrasts for F1 and F2 to identify vowels with >80% accuracy (Leek et al., 1987). Given that peak-to-valley contrasts of F2 for the three original vowels in this study were >6 dB, spectral contrasts of F2 of the original vowels were well resolved for NH listeners such that F2 enhancement did not benefit NH listeners to discriminate formant frequency difference.

On the other hand, the listeners with HI did show marked benefits (see the right panels of Figure 3) when F2 amplitude was enhanced at both 70 and 90 dB SPL, indicating the importance of spectral contrasts of formant peaks on vowel perception (Leek et al., 1987; Liu & Eddins, 2008). In order to examine audibility and presentation of the vowel F2 peaks for the listeners with HI, excitation patterns of the vowels /æ, i, ʌ/ with the original spectra and F2 enhancements at 70 and 90 dB SPL were computed following Moore and Glasberg's (2004) excitation pattern model. Excitation patterns were also calculated for the NH listeners.

In general, as shown in Figures 4 and 5, spectral peaks at high frequencies (e.g., F2 and F3) were smeared in the excitation patterns for the listeners with HI, likely due to the loss of frequency selectivity, whereas spectral peaks of F2 and F3 were well represented in the excitation patterns for the NH listeners (see Figure 6). As shown in Figure 4, a 9-dB F2 enhancement resulted in an enhanced representation of F2 in the excitation pattern for listeners with a mild-to-moderate hearing loss (e.g., S6). On the other hand, for listeners with a moderate-to-severe hearing loss (e.g., S1), F2 may not be audible, and the spectral peaks were severely smeared for the original vowels /i/ and /æ/ at 70 dB SPL, whereas F2 with the 9-dB enhancement may increase the F2 internal presentations and F2 audibility (see Figure 5). Thus, spectral enhancement of F2 peaks resulted in an enhanced internal representation and/or better audibility of F2 for the listeners with HI, especially those with severe hearing loss, thus improving their F2 discrimination.

However, vowel formant discrimination did not improve by simply increasing the overall speech level from 70 to 90 dB SPL with no F2 enhancement, suggesting that the difficulty of listeners with HI is accounted for not only by audibility, but also by reduced frequency selectivity (Needleman & Crandell, 1997). As indicated in Figures 4

Figure 4. Excitation pattern of vowels, /ʌ/ (top), /æ/ (middle), and /i/ (lower), with the original spectrum and a 9-dB enhanced F2 at 70 and 90 dB SPL for HI listener S6 (mild-to-moderate hearing loss). The audiogram of listener S6 was also plotted.

and 5, the internal F2 presentations became more smeared by an increase in speech level from 70 to 90 dB SPL for the listeners with HI, whereas the upward spread of

masking from F1 may increase as speech level increased, possibly accounting for the slightly worse vowel formant discrimination at 90 dB SPL. This is consistent with previous

Figure 5. Excitation pattern of vowels, /ʌ/ (top), /æ/ (middle), and /i/ (lower), with the original spectrum and a 9-dB enhanced F2 at 70 and 90 dB SPL for HI listener S1 (moderate-to-severe hearing loss). The audiogram of listener S1 was also plotted.

studies (i.e., only increasing speech level may not significantly compensate for spectral smearing of formant peaks for listeners with HI; Liu & Kewley-Port, 2007). However,

it should be noted that Coughlin et al. (1998) and Richie et al. (2003) reported that formant discrimination was improved from low (e.g., 60/70 dB SPL) to high (e.g., 95 dB

Figure 6. Excitation pattern of vowels, /ʌ/ (top), /æ/ (middle), and /i/ (lower), with the original spectrum and a 9-dB enhanced F2 at 70 and 90 dB SPL for the NH listeners.

SPL) levels. The differential effects of speech level on vowel formant discrimination among these studies could be due to the audibility and internal representations of F2 peaks. That is, in the studies of Coughlin et al. and

Richie et al., F2 was audible at high levels but not at low levels, and F2 was represented in the excitation patterns of their listeners with HI (for details, see Figure 6 of Liu & Kewley-Port, 2007).

In the present study, F2 audibility became better at the high speech level; however, F2 representation was severely smeared (see Figures 4 and 5), suggesting that increased audibility of F2 may improve formant discrimination only when F2 was represented well internally for listeners with HI. Moreover, in this study, HI listeners' performance improved by 14%–42% from 70 to 90 dB SPL as spectral enhancement of F2 was implemented, possibly due to better audibility of F2 peaks. For example, the F2 peak with a 9-dB enhancement was audible for the listener with moderate-to-severe hearing loss (e.g., S1; see Figure 5) at 90 dB SPL, but not at 70 dB SPL for vowels /i/ and /æ/. Altogether, these results suggest that both audibility and representation of formant peaks were critical to vowel perception for listeners with HI.

In the present study, the vowel stimuli with F2 enhancement were presented at the same level with the original vowels, resulting in an increase of the F2 spectral contrasts and a slight attenuation of the F1 amplitude (see Figures 4–6). Such acoustic changes may reduce the upward spread masking of F1 on the F2 and enhance the F2 representation (e.g., higher spectral contrasts of F2, resulting in better sensitivity to F2 frequency discrimination). This is consistent with the behavioral and physiological findings in other studies (Danaher et al., 1973; Danaher & Pickett, 1975; Miller et al., 1999a).

In the studies of F2 transition discrimination, listeners with HI performed better for stimuli with F1 removed than for stimuli with F1 present (Danaher et al., 1973; Danaher & Pickett, 1975). When the F2 and F3 peaks were amplified with no change in the F1 area, termed contrast-enhanced frequency shaping, the neural representation of the F2 frequency for the auditory nerve with noise-induced hearing loss was also enhanced (Miller et al., 1999a). The reduction of the F1 masking effect may also interpret individual variability regarding improvement of formant discrimination spectral contrasts among the listeners with HI. That is, the masking effect of F1 on F2 discrimination was greater for the listeners with a sloping high-frequency hearing loss compared to the listeners with a flat hearing loss such that the spectral enhancement of the F2 resulted in greater improvement in formant discrimination for the sloping high-frequency hearing loss (Danaher et al., 1973; Danaher & Pickett, 1975). For example, as shown in Table 1 and Figure 2, Listener S5 had a relatively flat hearing loss and showed less benefits than Listeners S1 and S4, who had a sloping high-frequency hearing loss. However, individual variability of listeners with HI may also be related to their hearing loss severity. For instance, Listeners S4 and S6 had similar hearing loss configurations, yet Listener S4 had greater improvement of formant discrimination than Listener S6 with relatively more severe hearing loss. More research is needed to investigate the effect of hearing loss configuration and severity on the spectral enhancement benefits of vowel formant discrimination.

The findings of this study are also consistent with behavioral and physiological studies in speech processing regarding spectral enhancement for the dysfunctional

auditory system (Bunnell, 1990; Lyzenga et al., 2002; Miller et al., 1999a; Stone & Moore, 1992). Bunnell (1990) manipulated spectral enhancement by applying filters in speech spectrum with different filter weights such that three sets of speech stimuli were produced in terms of spectral contrasts: normal, reduced, and enhanced. Results showed that stop consonant identification with spectral enhancement increased moderately (~5% in average) for listeners with HI compared to original speech sounds. Similarly, a 16-channel bandpass filter bank with each filter bandwidth close to the bandwidth of auditory filters of NH listeners was implemented by Stone and Moore (1992) to enhance spectral contrasts of speech sounds. Results of sentence recognition in noise indicated that there was no significant improvement in speech perception for the listeners with HI; however, listeners reported both higher quality and higher intelligibility of speech with spectral enhancement. Enhanced spectral contrasts of speech sounds combined with noise suppression also resulted in improvement of sentence recognition for listeners with simulated hearing loss (Lyzenga et al., 2002). Miller et al. (1999b) reported that the rate and temporal tonotopic representation of the F2 frequency in the abnormal auditory nerve was improved by amplifying the F2 spectral contrasts and F3 peaks.

Results of the present study indicate that spectral enhancement of vowel formant peaks may improve HI listeners' sensitivity to formant frequency changes. Given that thresholds of vowel formant discrimination were significantly related with vowel identification (e.g., lower thresholds of formant discrimination, better vowel identification; Kewley-Port, Bohn, & Nishi, 2005), improvement of vowel formant discrimination may help listeners with HI achieve greater accuracy in vowel identification. In addition, researchers have recently reported that vowels play a critical role in sentence recognition for young and elderly listeners with normal and impaired hearing (Fogerty, Humes, & Kewley-Port, 2010, 2012; Fogerty & Kewley-Port, 2009), implying that enhanced acoustic features of vowel sounds (e.g., spectral contrasts) may improve listeners' sentence recognition.

Overall, the present study and previous studies with spectral enhancement suggest that spectral enhancement of prominent peaks in speech spectrum may improve phonetic perception, speech recognition, or speech quality for listeners with HI. These findings could have implications for future hearing aid design as well as other assistive listening devices. On the other hand, the best performance of listeners with HI with a combined increase of speech level and spectral enhancement still could not reach the level of NH listeners, indicating that spectral enhancement may not be compensated solely by only an increase of the peak-to-valley contrasts, but perhaps also some other method, such as manipulation of formant bandwidth and/or spectral tilt. However, it should also be noted that spectral enhancement may result in changes of speech quality. More research is needed to study whether listeners with HI can benefit in vowel discrimination and identification from spectral enhancement.

Acknowledgments

A portion of the data was presented at the 158th Acoustic Society of America Meeting in San Antonio, TX. This study was funded by the University of Texas at Austin Undergraduate Research Fellowship to the first author. Special thanks are given to Craig Champlin for his comments on an earlier version of this manuscript.

References

- American National Standards Institute.** (2010). *Specification for audiometers* (Report No. S3.6-2010). New York, NY: Author.
- Bunnell, T. H.** (1990). On enhancement of spectral contrast in speech for hearing-impaired listeners. *Journal of the Acoustical Society of America*, 88, 2546–2556.
- Coughlin, M., Kewley-Port, D., & Humes, L. E.** (1998). The relation between identification and discrimination of vowels in young and elderly listeners. *Journal of the Acoustical Society of America*, 104, 3597–3607.
- Danaher, E. M., Osberger, M. J., & Pickett, J. M.** (1973). Discrimination of formant frequency transitions in synthetic vowels. *Journal of Speech and Hearing Research*, 16, 439–451.
- Danaher, E. M., & Pickett, J. M.** (1975). Some masking effects produced by low frequency vowel formants in persons with sensorineural hearing loss. *Journal of Speech and Hearing Research*, 18, 261–271.
- Fogerty, D., Humes, L. E., & Kewley-Port, D.** (2010). Auditory temporal-order processing of vowel sequences by young and older adults. *Journal of the Acoustical Society of America*, 127, 2509–2520.
- Fogerty, D., Humes, L. E., & Kewley-Port, D.** (2012). The relative importance of consonant and vowel segments to the recognition of words and sentences: Effects of age and hearing loss. *Journal of the Acoustical Society of America*, 132, 1667–1678.
- Fogerty, D., & Kewley-Port, D.** (2009). Perceptual contributions of the consonant-vowel boundary to sentence intelligibility. *Journal of the Acoustical Society of America*, 126, 847–857.
- Henry, B. A., Turner, C. W., & Behrens, A.** (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *Journal of the Acoustic Society of America*, 118, 1111–1121.
- Hillenbrand, J., Getty, L. J., Clark, M. J., & Wheeler, K.** (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099–3111.
- Humes, L. E., Dirks, D. D., Bell, T. S., & Kincaid, G. E.** (1987). Recognition of nonsense syllables by hearing-impaired listeners and by noise-masked normal hearers. *Journal of the Acoustical Society of America*, 81, 765–773.
- Kawahara, H., Masuda-Kastuse, I., & Cheveigne, A.** (1999). Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds. *Speech Communication*, 27, 187–207.
- Kewley-Port, D., Bohn, O.-S., & Nishi, K.** (2005). The influence of different native language systems on vowel discrimination and identification. *Journal of the Acoustical Society of America*, 117, 2399. Retrieved from <http://www.indiana.edu/~spl/l2vanfpdf.pdf>.
- Leek, M. R., Dorman, M. F., & Summerfield, Q.** (1987). Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 81, 148–154.
- Levitt, H.** (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49, 467–477.
- Liu, C.** (2008). Rollover effects of signal level on vowel formant discrimination. *Journal of the Acoustical Society of America*, 123, EL52–EL58.
- Liu, C., & Eddins, D. A.** (2008). Effects of spectral modulation filtering on vowel identification. *Journal of the Acoustical Society of America*, 124, 1704–1715.
- Liu, C., & Kewley-Port, D.** (2004). Vowel formant discrimination in high-fidelity speech. *Journal of the Acoustical Society of America*, 116, 1224–1233.
- Liu, C., & Kewley-Port, D.** (2007). Factors affecting vowel formant discrimination for hearing-impaired listeners. *Journal of the Acoustical Society of America*, 122, 2855–2864.
- Lyzena, J., Festen, J. M., & Houtgast, T.** (2002). A speech enhancement scheme incorporating spectral expansion evaluated with simulated loss of frequency selectivity. *Journal of the Acoustical Society of America*, 112, 1145–1157.
- Miller, R. L., Calhoun, B. M., & Young, E. D.** (1999a). Contrast enhancement improves the representation of /e/-like vowels in the hearing-impaired auditory nerve. *Journal of the Acoustical Society of America*, 106, 2693–2708.
- Miller, R. L., Calhoun, B. M., & Young, E. D.** (1999b). Discriminability of vowel representations in cat auditory-nerve fibers after acoustic trauma. *Journal of the Acoustical Society of America*, 105, 311–325.
- Moore, B. C. J., & Glasberg, B. R.** (2004). A revised model of loudness perception applied to cochlear hearing loss. *Hearing Research*, 188, 70–88.
- Mueller, H. G.** (2005). Fitting hearing aids to adults using prescriptive methods: An evidence based review of effectiveness. *Journal of the American Audiological Society*, 16, 448–460.
- Needleman, A. R., & Crandell, C. C.** (1997). Speech perception in noise by listeners with hearing impairment and simulated sensorineural hearing loss. In W. Jesteadt (Ed.), *Modeling sensorineural hearing loss* (pp. 461–473). Mahwah, NJ: Erlbaum.
- Richie, C., Kewley-Port, D., & Coughlin, M.** (2003). Discrimination and identification of vowels for hearing-impaired and masked normal-hearing listeners. *Journal of the Acoustical Society of America*, 114, 2923–2933.
- Stone, M. A., & Moore, B. C. J.** (1992). Spectral feature enhancement for people with sensorineural hearing impairment: Effects on speech intelligibility and quality. *Journal of Rehabilitation Research and Development*, 29, 39–56.
- Van Tasell, D. J.** (1980). Perception of second formant transitions by hearing-impaired persons. *Ear and Hearing*, 1, 130–136.

Copyright of American Journal of Audiology is the property of American Speech-Language-Hearing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.