Published in final edited form as:
J Speech Lang Hear Res. 2012 April ; 55(2): 474–486. doi:10.1044/1092-4388(2011/10-0238).

Consequences of broad auditory filters for identification of multichannel-compressed vowels

Pamela Souza,
Department of Communication Sciences and Disorders, Northwestern University

Richard Wright, and
Department of Linguistics, University of Washington

Stephanie Bor
Department of Speech and Hearing Sciences, University of Washington

Abstract

Purpose—In view of previous findings (Bor, Souza & Wright, 2008) that some listeners are more susceptible to spectral changes from multichannel compression (MCC) than others, this study addressed the extent to which differences in effects of MCC were related to differences in auditory filter width.

Method—Listeners were recruited in three groups: listeners with flat sensorineural loss, listeners with sloping sensorineural loss, and a control group of listeners with normal hearing. Individual auditory filter measurements were obtained at 500 and 2000 Hz. The filter widths were related to identification of vowels processed with 16-channel MCC and with a control (linear) condition.

Results—Listeners with flat loss had broader filters at 500 Hz but not at 2000 Hz, compared to listeners with sloping loss. Vowel identification was poorer for MCC compared to linear amplification. Listeners with flat loss made more errors than listeners with sloping loss, and there was a significant relationship between filter width and the effects of MCC.

Conclusions—Broadened auditory filters can reduce the ability to process amplitude-compressed vowel spectra. This suggests that individual frequency selectivity is one factor which influences benefit of MCC, when a high number of compression channels are used.

For many years, vowels were seen as "filler" between critical consonants that carried little information in themselves. Consider an example from popular culture: the voice of the "teacher" in animated Peanuts cartoons consisted of an unintelligible stream of vowel-like inflection, without any consonants (actually created with a trombone). More recently, researchers have focused on the contributions of the vowel to speech recognition (Kewley-Port, Burkle, & Lee, 2007). Segmentally, both the static center of the vowel and the transitory onset/offset boundaries carry useful information. For example, Rogers and Lopez (2008) found that speech recognition was impaired when the vowel center was removed, even though start and end formant transitions were intact. Vowels, then, deserve our attention for their contributions to speech recognition under different circumstances.

The most important identifying cues for a vowel are the frequency locations of the vowel formants (Hillenbrand, Houde, & Gayvert, 2006; Kewley-Port & Zheng, 1998; Klatt, 1982; Sommers & Kewley-Port, 1996; Syrdal & Gopal, 1986). Listeners can identify vowels given nothing but a sparse spectrum of tones at the formant frequencies (e.g., Molis, 2005). We can therefore view a vowel as a set of three to five spectral "peaks" separated by lower-amplitude "valleys". We will refer to the level difference between the peaks and valleys of the vowel spectra as spectral contrast.

To understand vowel identification, first consider representation of a vowel through a normal auditory system. With narrow auditory filters providing good frequency selectivity, the auditory system can easily resolve and distinguish the formant frequencies from the background, where the background can be either the spectral valleys of the vowel or a masking noise. A high level of identification is achieved especially when formant information is combined with other cues such as spectral shape and vowel duration.

Next consider the same vowel, represented through an impaired system in which auditory filters are broader than normal (Dubno & Dirks, 1989; Glasberg & Moore, 1986). In such a system, closely spaced formants may fall within the same critical band, so that the listener can no longer distinguish between vowels with similar formant frequencies. This is one reason why many listeners with sensorineural loss show degraded perception for naturally-produced vowels (e.g., Coughlin, Kewley-Port, & Humes, 1998; Liu and Kewley-Port, 2007; Richie, Kewley-Port and Coughlin, 2003).

Even if vowel formants are resolved into separate critical bands, widened auditory filters may smear spectral detail. Consequently, listeners with hearing loss require a larger spectral contrast (i.e., peak-to-valley ratio) to achieve the same performance as listeners with normal hearing when listening to vowel-like synthetic stimuli (Leek, Dorman and Summerfield, 1987; Turner & Holte 1987). The individuals tested in Leek et al. (1987) needed a peak-to-valley ratio of about 6 dB (in contrast to 2 dB for listeners with normal hearing). Because naturally-produced vowels have peak-to-valley ratios on the order of 5–30 dB, many vowels should still be identifiable despite the need for larger spectral contrast. However, accurate vowel identification will depend on the presented vowel, the listener's frequency selectivity, and whether alternative cues such as vowel duration are available.

The present study is concerned with effects of broadened auditory filters on vowel identification, *in combination with* effects of digital amplification. Digital hearing aids invariably apply some form of multichannel wide dynamic range compression (MCC). Briefly, in MCC the input signal is filtered into a number of channels. Gain is applied within each channel, with the goal of placing the output signal into the listener's dynamic range, and the channels are summed for the final output. Early researchers anticipated that a large number of compression channels might reduce spectral cues[1] (Lippman, Braida & Durlach, 1981; Plomp, 1988). Although there was little empirical testing of this effect, poorer performance with MCC was assumed to be due to reduction of spectral cues (Woods, Van Tasell, Rickert, & Trine, 2006). However, loss of spectral detail was also presumed to be of minor consequence for speech recognition when considered in view of the other benefits of MCC (Yund & Buckles, 1995).

Studies of MCC have focused mainly on consonant or sentence recognition, in which spectral, temporal, and contextual cues all contain usable information. Spectral degradation from MCC ought to have the greatest consequence for vowel identification, which depends on spectral cues. Bor, Souza, Wright (2008) undertook an acoustic analysis to quantify the acoustic and perceptual effects of MCC. Naturally-produced vowels were processed with one to 16 compression channels and a spectral contrast measure was developed to assess the effect of channel number. Most of the vowels showed significantly decreased spectral

[1]Compression amplification applies more gain for low-level inputs, and less gain for high-level inputs. With single-channel compression, the specified gain will be applied across the entire frequency range and spectral peaks and valleys will be maintained. With multi-channel compression, spectral valleys will receive more gain than spectral peaks. For a more detailed description of this effect, the reader is referred to Bor et al. (2008).

contrast, and reduced spectral contrast was associated with reduced vowel identification. It was of particular interest that recognition was reduced more for some listeners than others, even though they all had similar degrees of hearing loss.

To guide clinical choices, it would be useful to be able to determine which individuals might be most susceptible to such effects. It seems probable that the consequences of spectral change from MCC are magnified for individuals who have poorer representation of spectral detail due to broader auditory filters. Such a relationship has already been demonstrated, albeit for linearly-amplified consonants (Preminger and Wiley, 1985). Given that vowels depend more heavily on spectral cues than do consonants (Xu & Pfingst, 2008; Xu, Thompson, & Pfingst, 2005), we expect a relationship between auditory filter width and perception of MCC vowels.

The present study builds upon our previous work (Bor et al., 2008) to further investigate the effects of multichannel compression on vowel spectra, and in particular to explore reasons for individual variability. The following goals were addressed: first, to describe auditory filter characteristics for listeners with sloping and flat loss; second, to determine whether those groups showed a difference in identification of multichannel compressed vowels; and third, to determine whether performance differences across groups were related to auditory filter width.

Method

Listeners

Twenty-eight adults[2] participated in the study. The cohort included thirteen listeners with sloping loss (ages 53–88 years, mean age 74 years) and fifteen listeners with flat loss (ages 44–88 years, mean age 71 years). Listeners were tested monaurally and a loss that met the criteria in one ear was sufficient for inclusion. All listeners had sensorineural hearing loss, defined as no air-bone gap greater than 10 dB at octave frequencies from .25 to 4 kHz and static admittance and tympanometric peak pressure within normal limits in the test ear according to Wiley et al. (1996). Each listener with sloping loss had normal hearing or a mild loss (≤35 dB HL[3]) at 500 Hz; with 2000 Hz thresholds at least 20 dB worse than 500 Hz thresholds. Each listener with flat loss had a moderate loss (40–60 dB HL) at 500 Hz; with 500 and 2000 Hz thresholds within 10 dB of each other. Mean hearing thresholds for the two groups are shown in Table 1. Our primary interest was not audiometric configuration *per se*, but rather the relationship between auditory filter width and identification of MCC vowels. The choice of listeners with flat and sloping loss was a convenient way to increase the probability that auditory filter patterns would vary across groups. Because pure-tone threshold and auditory filter width are moderately correlated (Carney & Nelson, 1983; Dubno & Dirks, 1989; Florentine, Buus, Scharf, & Zwicker, 1980; Stelmachowicz, Jesteadt, Gorga, & Mott, 1985), we expected that (1) low-frequency auditory filters would be wider for the listeners with flat than sloping loss; (2) high-frequency auditory filters would be similar across the two groups; and (3) there would be within-group variability.

There was no significant difference in age between groups (t_{26}=−.68, p=.506). Each listener completed the Mini-Mental State Exam and had at least the minimum score of 26 (out of 30) considered to represent normal cognitive function (Folstein, Folstein, & McHugh, 1975). Group mean MMSE scores were 29.5 points for the sloping group and 29.0 points for the flat group, with no significant difference between groups (t_{26}=1.59, p=.454).

[2]Number of subjects was based on 80% power to detect the main effect of interest (Bausell & Li, 2002).
[3]Threshold values are expressed in dB HL re: ANSI, 2004.

Data were also collected for four listeners with normal hearing (ages 26–33 years, mean age of 28 years). All of the listeners with normal hearing had pure-tone thresholds of 20 dB HL or better at octave frequencies from .25 to 8 kHz. The normal-hearing data were used as a reference point for data interpretation.

English was the first or primary language for all listeners. All procedures were reviewed and approved by the local Institutional Review Board and listeners were reimbursed for their time.

Auditory filter characteristics

Auditory filter characteristics were estimated with an adaptation of the procedure used by Leek and Summers (1996). Masked thresholds were determined for sinusoidal tones in the presence of notched-noise maskers. Tone frequencies were .5 or 2 kHz, based on the auditory filters expected to be critical to the vowel identification task (F1 mean of .53 kHz and F2 mean of 1.9 kHz across all vowels).

Stimuli

All signals were digitally generated in Matlab at a 48.8 kHz sampling rate. The target tones were 360 msec in duration (including 25 ms cosine-squared rise and fall ramps). The tones were presented at 10 dB above the listener's hearing threshold at the tone frequency.[4]

The masker was 460 ms in duration (including 25 ms cosine-squared rise and fall ramps). When masker and tone were presented together, the tone was temporally centered within the masker (i.e., the masker began 50 ms before and ended 50 ms after the tone). The masker was digitally generated in the frequency domain, resulting in two bands of noise located to either side of the target tone frequency. The maskers were placed either symmetrically or asymmetrically around the tone frequency, creating a variety of notch widths (Table 2). Notch widths were chosen based upon recommended (abbreviated) parameters for estimating auditory filter widths (Stone, Glasberg, & Moore, 1992). The notch width was defined as the deviation of each edge of the central notch from the center frequency (f_c), and was denoted as $\Delta f / f_c$. There were six notch widths: symmetrical widths of 0.0, 0.1, 0.2 and 0.4; plus two asymmetrical widths of 0.2 and 0.4 for the low and high band, and 0.4 and 0.2 for the low and high bands. The lower and upper edges for the notched noise were set at 0.8 times f_c (Baker and Rosen, 2002). Thus, the outer edges of the noise were 100 and 900 Hz for the 500 Hz tone, and 400 and 3600 Hz for the 2000 Hz tone.

Procedure

At each notch width, the tone level was fixed and the masker level was varied adaptively to obtain thresholds. The tone and masker were mixed and presented monaurally to the test ear via an ER2 insert earphone. The masked threshold was estimated for each notch width using a two-alternative forced-choice paradigm with feedback. In each trial, the masker was present in both intervals and the tone was present in one of those intervals. Listeners were instructed to select the interval with the tone, using a touch screen. In the first trial, signal-to-noise ratio was 0 dB. The masker increased in 5 dB steps after three consecutive correct responses and decreased in 5 dB steps after one incorrect response. The 5 dB step size decreased to 2 dB after four reversals and continued until 10 additional reversals occurred, at which point the threshold was calculated as the mean tone level of the last six reversals. A single threshold measurement was referred to as a "block".

[4]The presentation level was increased to 13 dB SL for three listeners because the level of the noise could not be lowered far enough for those listeners to detect the tone.

To familiarize the listener with the task, threshold was measured for each listener with the tone presented at 70 dB SPL using a symmetrical notch width of 0.2. Next, the probe level was fixed at 10 dB SL (re: audiometric threshold). For each listener, twenty-four thresholds (6 notch widths x 2 tone frequencies x 2 test blocks) were measured. Presentation order of the notch width and tone frequency was randomized.

Estimating the auditory filter

Auditory filter width was estimated using a *roex* (*p*, *r*) model, where *p* was the width of the filter's passband and *r* was the dynamic range (Patterson, Nimmo-Smith, Weber, & Milroy, 1982). The Polyfit program (Rosen, Baker & Darling, 1998) was used to calculate the auditory filter parameters and final ERB values for this study. The model was specified to allow three polynomial terms for both the lower and upper slope of the filter, and the allowed shift for off-place listening was set at 0.2.

Unlike Leek and Summers (1996), we opted not to average discrepant data into the final value but rather to discard those points that according to a quality check did not represent a valid response to the stimuli. A data point was considered invalid if standard deviation within a block exceeded 5 dB. When invalid data points were noted during testing the same condition was repeated in an attempt to obtain valid data. Ninety-seven percent of the data points passed the quality check and were retained. For those data, the between-block difference was less than 1 dB averaged across all listeners, test tones and notch widths. Correlation between the values for the first and second block was r =0.89. In those cases, results for block 1 and block 2 were averaged together for the final notch width result. Otherwise, results from a single block were used.

Although all listeners were able to complete the task, some estimated filters could not be fit. A modeled filter slope of zero indicated a poorly fit function. In the case of filter slope zero, each notched noise condition was systematically removed from and/or added to the calculation to better model the filter function (M. Leek, personal communication). In some cases it was not possible to avoid a filter slope of zero by adding or subtracting conditions. The filter could not be modeled at .5 kHz for two listeners with flat loss and one listener with sloping loss and at 2 kHz for six listeners with flat loss and three listeners with sloping loss. Dubno and Dirks (1989) noted a similar issue, in which masked-threshold data for some listeners could not be appropriately fit by the two-parameter roex model, despite those listeners having similar hearing thresholds as the rest of the listener cohort.

Vowel Identification

Stimuli

Stimuli consisted of naturally-produced vowels by six adult male and six adult female talkers. A detailed description of the vowel recording and selection can be found in Bor et al. (2008). Briefly, all vowels were recorded in a /hVd/ or /Vd/ context, including "heed" / i/, "hid" / ɪ/, "aid" /e/, "head" /ε/, "had" /æ /, "odd" /ɑ/, "hood" /ʊ/, "who'd" /u/. The eight words were randomized and repeated five times and words from equivalent positions in the set of repetitions were used to control for list effects on pronunciation. All talkers were instructed to read the list of words at a natural pace and vocal intensity. Vowels were recorded at 44.01 kHz or 22.05 kHz and downsampled to 11.025 kHz.

One representative token of each vowel was selected for each talker based on recording fidelity, clarity of each talker's voice, and similarity across vowels in the intonation. Linear predictive coding (LPC) formant tracks overlaid on a wideband spectrogram and pitch tracks overlaid on a narrowband spectrogram were used to ensure formant and pitch steady states at the vowel midpoint. The final set consisted of 96 words (12 talkers x 8 vowels). Table 3 lists mean and standard error of the first and second formant frequencies for the vowel set. Peak amplitude was normalized across vowels. For these steady-state signals, final RMS levels across vowels varied by less than 2 dB.

Compression

To create the MCC vowels, the 96 /hVd/ words were digitally filtered into sixteen channels using 5-pole, Butterworth 1/3rd-octave band filters. The lower-to-upper range across all channels was set at 141 Hz and 5,623 Hz. Division points between channels were based on equal-octave spacing (Table 4). The channels were compressed separately using simulation software then summed together to create the final MCC signals.

The compression parameters were the same in each channel and included a compression ratio of 3:1, attack time of 3 ms and and release time of 50 ms. Compression threshold was set at 30 dB below the peak level of the normalized vowel. Note that parameters suitable to test the experimental question took priority. For that reason, the compression parameters were deliberately not varied across individuals (or across channels) as would be done in the clinic. However, the compression parameters were within the range of those in clinical use. Output level of the vowels was dictated by the application of an individual frequency-gain response, described below.

To remove formant transitions and vowel duration cues a 150 ms segment was extracted from the center of each vowel. The 150 ms duration was chosen after examining vowel duration among the entire set. Because all vowels were longer than 150 ms, the brief temporal overshoot of the attack time was removed. A 5 ms linear amplitude ramp was applied to the onset and offset of each segment.

A linear (uncompressed) condition was created using all process steps described here except that the stimuli were bandpass filtered from 141 Hz to 5,623 Hz and there was no compression simulator.

In order to ensure sufficient audibility of all vowels for the listeners with hearing loss and to mimic a clinical scenario, amplification was added to the vowels prior to presentation. An individual frequency response was calculated for each listener using NAL-RP prescriptive values (Byrne & Dillon, 1986) and used to create output targets (output as function of frequency) that were the same for the linear and MCC conditions. Output levels as a function of frequency were measured for each listener to ensure that targets were met, and that there were no differences in audibility across conditions. An additional random attenuation of 1 or 2 dB was applied to prevent any unanticipated intensity biases.

Procedure

Listeners were seated in a double-walled sound booth. To familiarize the listener with the orthography of the vowel choices, each listener completed a training task at least twice, or until performance was at least 88% correct. The task was to match the orthographic representation of the vowel sound to a set of three words that had the same vowel sound, using an eight-alternative forced-choice procedure. For example, /i/ was orthographically represented as “ee” and example words were “cheese”, “meat” and “leaf”. There was no auditory signal during this task. Correct-answer feedback was given.

To measure vowel identification, stimuli were presented monaurally to the listeners via an ER3 insert earphone. The test ear for each listener was the same as the test ear for the notched-noise experiment.

A block consisted of 96 randomly ordered trials (8 vowels x 12 talkers). To account for learning effects (discussed in detail below), listeners completed four blocks in each condition (compressed or linear). Half of the listeners heard the compressed blocks first, and half heard the linear blocks first. Listeners responded to each trial in an 8-alternative forced-choice paradigm using a touch screen. The location of response buttons on the touch screen was randomized for each block to prevent response bias.

Results

Auditory filter width

Figure 1 shows the auditory filter equivalent rectangular bandwidth (ERB) for each listener group as a function of audiometric threshold. Larger ERB values indicate broader filters. Listeners with poorer thresholds tended to have wider auditory filters, but there was considerable variability.

At 500 Hz, there was a significant difference among groups ($F_{2,26}$=4.84, p=.017). Post-hoc analyses (Fisher's LSD) indicated that low-frequency (.5 kHz) filters were wider for the listeners with flat loss than for listeners with normal hearing (p=.007) and marginally wider for listeners with flat loss than those with sloping hearing loss (p=.056). There was no significant difference between listeners with sloping loss and with normal hearing (p=.137).

At 2000 Hz, there was a significant difference among groups ($F_{2,26}$=5.57, p=.011). Between–group post-hoc comparisons showed no significant difference between listeners with flat and sloping loss (p=.941). The listeners with normal hearing had narrower filters than both the flat (p=.008) and sloping groups (p=.004).

ERB values for 500 and 2000 Hz were correlated with each other (r=.42, p=.017). The significant but modest correlation reflects that although listeners with wide low-frequency filters tend to have wide high-frequency filters, having a wide-high frequency filter does not guarantee a widened low-frequency filter. That finding was consistent with a cohort that included listeners with sloping audiograms, for whom we would expect narrow low- and wide high-frequency filters.

Vowel Identification

Learning on the task was minimal, with very similar scores across blocks. Within each condition (linear and MCC), mean scores improved by less than 5 percentage points on subsequent blocks, and performance stabilized (defined as a non-significant difference in scores, p>.05) between blocks two and three. Accordingly, scores for blocks two, three and four were averaged and used for all subsequent data analysis.

Scores for each group and amplification condition are summarized in Figure 2. There are several notable effects. First, the group mean performance is always poorer for the MCC than for the linear condition, regardless of hearing status. In our paradigm, frequency-gain responses (and output levels) were the same for the linear and MCC vowels for each listener. Therefore, MCC did not provide any audibility or loudness comfort advantage, and the differences in scores between amplification conditions reflect the MCC processing.

Second, performance was better for listeners with sloping than with flat loss. Because both groups received appropriate audibility and were of similar age and cognitive status, it is likely that this effect is due to differences in cochlear processing. It is also notable that variability in the sloping loss group was quite small for linear amplification, but increased with MCC. This suggests that listeners with sloping loss with similar audiograms and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

similar recognition of linearly-processed speech may respond differently to processed signals.

The percent correct scores were converted to rationalized arcsine units (RAUs; Studebaker, 1985) and a two-way repeated-measures ANOVA was completed to compare across hearing loss group and amplification condition. Identification scores were worse for MCC than linear speech ($F_{1,29}$=40.94, p<.005) and for listeners with flat loss than listeners with sloping loss ($F_{2,29}$=13.19, p<.005). There was no interaction between hearing loss group and amplification condition ($F_{2,29}$=2.01, p=.153).

Relationship between auditory filter width and response to MCC

The analysis described above indicated that MCC had a similar (negative) effect on vowel identification for listeners with flat and sloping loss. However, the specific hypotheses motivating this work were that (1) frequency selectivity would vary within a hearing loss group and that (2) response to MCC would be influenced by frequency selectivity. The first hypothesis was supported by the data in Figure 1, where not all listeners with a given audiometric configuration had similar frequency selectivity. The second hypothesis was tested with the following analysis.

First, the ERB values were converted to relative values by dividing each ERB (in Hz) by the test frequency. For example, an ERB of 200 Hz at a center frequency of 500 Hz converts to a relative ERB of 0.4, and an ERB of 200 Hz at a center frequency of 2000 Hz converts to a relative ERB of 0.1. Next, the relative ERB values at 500 and 2000 Hz were averaged to produce a single per-listener value that (roughly) represented the listener’s frequency selectivity. Figure 3 shows vowel recognition as a function of average relative ERB for linear (filled circles) and MCC (open triangles) vowels. Best-fit lines are shown for linear (solid line) and MCC (dashed line). Although there is considerable variability, particularly for the MCC vowels, scores decreased at higher average relative ERBs. A multiple-regression analysis was used to determine the amount of variance in vowel identification that was accounted for by average relative ERB and type of amplification. In designing the regression model, we expected MCC to interact with auditory filter width. Consider an example: the vowel /ɑ/ spoken by a female talker with F1 about 800 Hz and F2 about 1200 Hz. In the compression algorithm used here, those formants would fall into channels 8 and 10 (Table 4). Even with reduced spectral contrast from MCC, a listener with narrow filter width may be able to resolve those closely spaced formants and correctly identify the vowel. However, for a listener with impaired frequency selectivity, the combination of reduced spectral contrast from MCC coupled with reduced spectral representation of the vowel due to loss of auditory tuning may result in an error. Accordingly, an interaction term representing the interaction between ERB and amplification type was also included in the model. Contributions of the main-effect variables and the interaction were examined in a stepwise-regression procedure, using alpha-level criteria of .05 for probability of entry into the model and .10 for probability of removal from the model. The final model accounted for 85% of the variance in the score ($F_{3,63}$=113.57, p<.005). Significant predictors were amplification type (β=−.16, t=.286, p<.005), average ERB (β =−1.18, t=−15.19, p<.005) and the ERB by amplification interaction (β =1.07, t=13.08, p<.005). In other words, vowel identification is better for linear than MCC vowels; and is better when the listener has good frequency selectivity. The significant interaction (illustrated by the solid and dashed fit lines in Figure 3) indicates that listeners with good frequency selectivity, who would otherwise be able to resolve formant information, are more detrimentally affected by MCC than listeners with naturally-poorer frequency selectivity.

Although not an experimental hypothesis, it is also of interest to examine the relationship between response to MCC and pure-tone threshold. Similar to the treatment of the relative

ERB values, pure-tone thresholds (dB HL) at 500 and 2000 Hz were averaged to produce a single per-listener value that (roughly) represented the listener's hearing status. Table 5 shows pairwise correlation values (Pearson *r*) for average hearing threshold, average relative ERB and vowel identification.

Vowel error patterns

Figure 4 shows the vowel error patterns, plotted as graphical confusion matrices. In each panel, presented (target) vowels (ordered by F1 frequency) are shown on the abscissa and responses are shown on the ordinate. The size of each circle represents the number of responses for a particular presented vowel/responded vowel combination. Responses on the diagonal are correct, and those are noted by percent correct values. Responses off the diagonal are incorrect. For example, for listeners with sloping loss listening through linear amplification, the vowel /i/ was correctly identified 96% of the time, and errors were about evenly distributed among the other vowels, with the exception of /e/ which was never chosen. For listeners with sloping loss listening through MCC amplification, the vowel /i/ was identified correctly only 62% of the time, and the most common errors were /I/ and /Ɛ/. Qualitatively, the greater the "scatter" to either side of the diagonal, the more difficult the task. Mirroring the data from Figure 2, the largest number of errors occurred for listeners with flat loss presented with MCC vowels, and the smallest number of errors occurred for listeners with sloping loss presented with linear vowels.

Discussion

Considerations when measuring auditory filter width

Listeners with flat versus sloping loss are assumed to have different frequency resolution, particularly at low frequencies. In our data, listeners with flat loss had broader auditory filters than listeners with sloping loss at low frequencies, and similar filter widths at high frequencies. The exact width of the auditory filter varied among individuals even when they had similar hearing thresholds. Numerous investigators have demonstrated variability in frequency resolution (auditory filter width) among individuals (e.g., Carney & Nelson, 1983; Hopkins & Moore, 2011; Florentine et al., 1980; Stelmachowicz et al., 1985; Turner & Henn, 1989). Indeed, the pure-tone audiogram is now seen as a relatively coarse measure of hearing function, as demonstrated by recent calls for more sensitive hearing assessment measures (Abel, Siegel, Banakis, Chan, Zecker, & Dhar, 2009; Souza & Tremblay, 2006; Walden & Walden, 2004).

For the present data, average hearing threshold was at least as strongly related to vowel performance as the more specific filter measurements (Table 5). Earlier work by Turner and Henn (1989) suggests that the hearing threshold-to-performance correlation is less related to audibility *per se* than to the relationship between auditory threshold and frequency selectivity. Listeners with better hearing thresholds tended to be those with good frequency selectivity and they were also the listeners with the best vowel identification.

There are some practical problems with estimating auditory filters using the method reported here. Some listeners, particularly those with poorer thresholds, have auditory filter bandwidths that are too broad for precise mathematical modeling. In our data, more of these listeners were noted in the flat loss group and were excluded from data analysis. Some modeling errors can be avoided if test conditions include more repetitions at each notch-width, or if more notch-width conditions are included, but the amount of test time is already substantial. The average time required to obtain data for just two auditory filters was two hours per listener, rendering such testing inappropriate for clinical use. We have begun evaluating some “fast” auditory filter measurements that could be adapted for the clinic (Charaziak, Souza, & Siegel, submitted).

A related issue is choosing which auditory filters to measure. We selected auditory filters centered at 500 and 2000 Hz as representative of the frequency region for our vowel formants. Those frequencies were a reasonable, albeit sparse, description of across-frequency selectivity for each listener, but are only a part of the abilities needed to perform well on the vowel task as a whole. We did not necessarily measure the auditory filter the listener would need to identify a specific formant for a specific vowel by a specific speaker. Also, identifying a formant is essential, but not sufficient, for vowel identification. The listener also needs to “reconstruct” the vowel across frequency; that is, to identify formants relative to other formants. Apoux and Healy (2009) estimated that 20 auditory filters are required for accurate vowel identification. In that sense, no single auditory filter measurement can predict vowel identification. Nonetheless, it is likely that listeners with broadened .5 and 2 kHz filters also have broadened filters at other frequencies, and that listeners with poor frequency selectivity will have difficulty perceiving speech elements that rely on spectral contrast, such as vowels.

Influence of auditory filter width on identification of MCC vowels

Vowel identification worsened with increasing auditory filter width, and when multichannel compression was used. There is evidence that both manipulations smooth the vowel spectra received by the listener. With regard to internal spectral representation, Leek and Summers (1996) tested listeners with hearing loss to determine how much spectral contrast was needed before they were able to separate out target components from background components. They assumed that the detectable difference in intensity between target peaks and background harmonics indicated how much contrast was preserved in the cochlear excitation pattern. Listeners with broader high-frequency auditory filters were less able to use F2 frequency differences to make identification judgments and based their decision almost entirely on F1 peaks. Several authors (Leek & Summers, 1996; Turner & Henn, 1989) suggested that broader auditory filters create an altered internal representation (excitation pattern) of a vowel. Effectively, the output of each auditory filter results in a poorer signal-to-noise ratio than would be the case for a listener with a more sharply tuned auditory system. Animal models also support loss of frequency selectivity in a damaged cochlea (Miller, Calhoun & Young, 1999).

Bor et al. (2008) demonstrated that increasing number of compression channels causes greater alteration to vowel spectra. They argued that overuse of MCC (i.e., using a large number of compression channels coupled with moderate to high compression ratios) could smooth vowel spectra in a manner akin to that of broadened auditory filters. After decades in which hearing aids were marketed with ever-larger numbers of compression channels, there is growing acknowledgment that more is not always better, or at least that the level of acoustic manipulation should be customized for the listener (e.g., Edwards, 2007; Souza, 2011; Woods et al. 2006). The interaction between auditory filter width and amplification type seen in the present data suggests that the spectral smoothing resulting from MCC has a relatively greater effect on a listener with narrower auditory filters, who would otherwise be able to resolve vowel formant peaks.

When multichannel compression is needed for other reasons, such as loudness comfort or audibility across frequency, it might be possible to counteract the effects of spectral smoothing with additional processing. Spectral sharpening has been suggested for use in hearing devices (DiGiovanni, Nelson, & Schlauch, 2005; Oxenham, Simonson, Turicchia, & Sarpeshkar, 2007). However, Franck, van Kreveld-Bos, Dreschler and Verschuure (1999) noted that vowel perception was poorer when 8-channel MCC was combined with spectral

NIH-PA Author Manuscript

NIH-PA Author Manuscript

sharpening, than with spectral sharpening alone. In contrast, combining single-channel compression with spectral sharpening had no effect compared to spectral sharpening alone. Although MCC alone (without sharpening) was not tested, the negative effects of MCC demonstrated by Franck et al. were consistent with the present data.

Relating current data to clinical use of MCC

To what extent will the data presented here generalize to wearable hearing aids in everyday listening situations? We know that listeners with hearing loss rely more heavily on vowel spectra than on dynamic formant transitions (e.g., Lentz, 2006), so the signals used in this study should have captured the effect of distortion on essential acoustic cues. However, our signals were more strictly controlled than would be the case in everyday listening. Under more realistic conditions, vowels amplified through hearing aids would contain other cues such as vowel duration and formant transitions, and would also have lexical and contextual information from the word or phrase in which the vowel is embedded. We deliberately controlled for temporal confounds, such as vowel duration. The post-compression processing used here also eliminated any attack time overshoots, thus the data should be relevant to either slow- or fast-acting provided the compression threshold is sufficiently low that compression is engaged throughout the duration of the vowel. In a very slow-acting MCC system where some or all of the vowel is linearly amplified, performance might be more similar to the linear vowel scores presented here. Finally, while the spectrally flattened vowels used in this study were created with MCC parameters that were clinically feasible, an individual listener might receive more or less spectral contrast than represented here depending on the prescriptive procedure, compression parameters, crossover frequencies, signal input level, and presence or absence of background noise. Crain and Yund (1995) noted that MCC affected vowel identification with some, but not all, compression parameters and recommended that individual adjustments be made for each listener.

The present study compared MCC and linearly-amplified speech, where the control (linear) condition was more similar to a single-channel aid with frequency shaping than to a digital hearing aid designed to function as a multichannel compressor but set to linear (i.e., compression ratio of 1:1 in each compression channel). Specifically, a digital hearing aid set to linear would filter the speech into channels (without compression) then mix the output of the filters. Given the experimental focus on spectral cues, we chose a Butterworth filter which minimized spectral distortion within and outside the passband. Such a filter can introduce a delay in the temporal domain, particularly for very narrow filters. Spectral analysis of our filtered signals pre-and post-compression indicates that such distortion was minimal, in part because there was little energy present in the input signal for the narrowest (lowest-frequency) filters. Moreover, even much greater amounts of temporal distortion are not expected to impact recognition of stationary-formant vowels (see Assmann & Summerfield, 2005, p. 270, for discussion of this issue).

In summary, hearing technology has become increasingly complex. First-generation digital hearing aids had no more than two or three channels, while today's MCC hearing aids provide as many as 20 channels. More channels have been viewed as a positive aspect of processing in terms of providing better loudness comfort, precise gain adjustments, and the potential for better noise control and feedback reduction (Chung, 2004a; b). However, beyond some therapeutic point, there may be diminishing returns. Audibility can be maximized with as few as four channels (Woods et al., 2006). Despite a longstanding industry trend towards more sophisticated processing, the most appropriate solution for some listeners may be simpler processing that preserves signal features.

Acknowledgments

The authors thank Kathryn Arehart, Marjorie Leek, Michelle Molis, and Chris Stecker for sharing code and expertise for portions of this project, and Star Reed and James Dewey for help with data analysis. Andy Sabin and Jon Siegel provided helpful suggestions regarding filter measurements. We also thank Steve Armstrong and Gennum Corporation for providing the software used to simulate hearing aid compression. Work was supported by NIH grant R01 DC0060014.

References

Abel, R.; Siegel, J.; Banakis, R.; Chan, CL.; Zecker, S.; Dhar, S. Behavioral audiograms measured in humans using a multipurpose instrument. Presented at theMidwinter Meeting of the Association for Research in Otolaryngology; Baltimore, MD. 2009.

ANSI. American National Standards Institute Specification for Audiometers (Vol. S3.6-2004). 2004

Apoux F, Healy EW. On the number of auditory filter outputs needed to understand speech: further evidence for auditory channel independence. Hearing Research. 2009; 255(1–2):99–108. [PubMed: 19539016]

Baker RJ, Rosen S. Auditory filter nonlinearity in mild/moderate hearing impairment. Journal of the Acoustical Society of AmericaJournal of the Acoustical Society of America. 2002; 111(3):1330–1339.

Bausell, RB.; Li, Y-F. Power Analysis for Experimental Research: A Practical Guide for the Biological, Medical and Social Sciences. New York: Cambridge University Press; 2002.

Bor S, Souza P, Wright R. Multichannel compression: effects of reduced spectral contrast on vowel identification. Journal of Speech Language and Hearing Research. 2008; 51(5):1315–1327.

Byrne D, Dillon H. The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. Ear and Hearing. 1986; 7:257–265. [PubMed: 3743918]

Carney AE, Nelson DA. An analysis of psychophysical tuning curves in normal and pathological ears. Journal of the Acoustical Society of AmericaJournal of the Acoustical Society of America. 1983; 73(1):268–278.

Charaziak K, Souza P, Siegel J. Time-efficient measures of auditory frequency selectivity. submitted.

Chung K. Challenges and recent developments in hearing aids. Part I. Speech understanding in noise, microphone technologies and noise reduction algorithms. Trends in Ampification. 2004a; 8(3):83–124.

Chung K. Challenges and recent developments in hearing aids. Part II. Feedback and occlusion effect reduction strategies, laser shell manufacturing processes, and other signal processing technologies. Trends in Amplification. 2004b; 8(4):125–64. [PubMed: 15735871]

Coughlin M, Kewley-Port D, Humes LE. The relation between identification and discrimination of vowels in young and elderly listeners. Journal of the Acoustical Society of AmericaJournal of the Acoustical Society of America. 1998; 104(6):3597–3607.

Crain TR, Yund EW. The effect of multichannel compression on vowel and stop-consonant discrimination in normal-hearing and hearing-impaired subjects. Ear and Hearing. 1995; 16(5): 529–43. [PubMed: 8654908]

DiGiovanni JJ, Nelson PB, Schlauch RS. A psychophysical evaluation of spectral enhancement. Journal of Speech Language and Hearing Research. 2005; 48(5):1121–35.

Dubno JR, Dirks DD. Auditory filter characteristics and consonant recognition for hearing-impaired listeners. Journal of the Acoustical Society of America. 1989; 85(4):1666–1675. [PubMed: 2708682]

Edwards B. The future of hearing aid technology. Trends in Amplification. 2007; 11:31–45. [PubMed: 17301336]

Florentine M, Buus S, Scharf B, Zwicker E. Frequency selectivity in normally-hearing and hearing-impaired observers. Journal of Speech and Hearing ResearchJournal of Speech Language and Hearing Research. 1980; 23(3):646–669.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; 12(3):189–198. [PubMed: 1202204]

Franck BA, van Kreveld-Bos CS, Dreschler WA, Verschuure H. Evaluation of spectral enhancement in hearing aids, combined with phonemic compression. Journal of the Acoustical Society of AmericaJournal of the Acoustical Society of America. 1999; 106(3 Pt 1):1452–64.

Glasberg BR, Moore BC. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. Journal of the Acoustical Society of AmericaJournal of the Acoustical Society of America. 1986; 79(4):1020–1033.

Hillenbrand JM, Houde RA, Gayvert RT. Speech perception based on spectral peaks versus spectral shape. Journal of the Acoustical Society of AmericaJournal of the Acoustical Society of America. 2006; 119(6):4041–4054.

Hopkins K, Moore BCJ. The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. Journal of the Acoustical Society of America. 2011; 130:334–349. [PubMed: 21786903]

Kewley-Port D, Zheng Y. Auditory models of formant frequency discrimination for isolated vowels. Journal of the Acoustical Society of America. 1998; 103(3):1654–1666. [PubMed: 9514029]

Kewley-Port D, Burkle TZ, Lee JH. Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. Journal of the Acoustical Society of America. 2007; 122(4):2365–2375. [PubMed: 17902871]

Klatt DH. Prediction of perceived phonetic distance from critical-band spectra: A first step. Proceedings of the IEEE International Conference on Speech Acoustics and Signal Processing. 1982; 129:1278–1281.

Leek MR, Dorman MF, Summerfield Q. Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America. 1987; 81:148–154. [PubMed: 3819173]

Leek MR, Summers. Reduced frequency selectivity and the preservation of spectral contrast in noise. Journal of the Acoustical Society of America. 1996; 100(3):1796–806. [PubMed: 8817905]

Lentz JJ. Spectral-peak selection in spectral-shape discrimination by normal-hearing and hearing-impaired listeners. Journal of the Acoustical Society of America. 2006; 120(2):945–956. [PubMed: 16938982]

Lippmann RP, Braida LD, Durlach NI. Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss. Journal of the Acoustical Society of America. 1981; 69:524–534. [PubMed: 7462475]

Liu C, Kewley-Port D. Factors affecting vowel formant discrimination by hearing-impaired listeners. Journal of the Acoustical Society of America. 2007; 122(5):2855–2864. [PubMed: 18189575]

Miller RL, Calhoun BB, Young ED. Discriminability of vowel representations in cat auditory-nerve fibers after acoustic trauma. Journal of the Acoustical Society of America. 1999; 105:311–325. [PubMed: 9921658]

Molis M. Evaluating models of vowel perception. Journal of the Acoustical Society of America. 2005; 118(2):1062–71. [PubMed: 16158661]

Oxenham AJ, Simonson AM, Turicchia L, Sarpeshkar R. Evaluation of companding-based spectral enhancement using simulated cochlear-implant processing. Journal of the Acoustical Society of America. 2007; 121(3):1709–16. [PubMed: 17407907]

Patterson RD, Nimmo-Smith I, Weber DL, Milroy R. The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold. Journal of the Acoustical Society of America. 1982; 72(6):1788–1803. [PubMed: 7153426]

Plomp R. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. Journal of the Acoustical Society of America. 1988; 83(6):2322–2327. [PubMed: 3411024]

Preminger J, Wiley TL. Frequency selectivity and consonant intelligibility in sensorineural hearing loss. Journal of Speech and Hearing Research. 1985; 28(2):197–206. [PubMed: 4010249]

Richie C, Kewley-Port D, Coughlin M. Discrimination and identification of vowels by young, hearing-impaired adults. Journal of the Acoustical Society of America. 2003; 114(5):2923–2933. [PubMed: 14650026]

Rogers CL, Lopez AS. Perception of silent-center syllables by native and non-native English speakers. Journal of the Acoustical Society of America. 2008; 124(2):1278–93. [PubMed: 18681614]

Rosen S, Baker RJ, Darling A. Auditory filter nonlinearity at 2 kHz in normal hearing listeners. Journal of the Acoustical Society of America. 1998; 103(5 Pt 1):2539–2550. [PubMed: 9604348]

Sommers MS, Kewley-Port D. Modeling formant frequency discrimination of female vowels. Journal of the Acoustical Society of America. 1996; 99(6):3770–3781. [PubMed: 8655808]

Souza, P. Top-down and bottom-up hearing aid processing. Paper presented at the American Academy of Audiology meeting; Chicago, IL. 2011.

Souza P, Tremblay K. New perspectives on assessing amplification effects. Trends in Amplification. 2006; 10(3):119–144. [PubMed: 16959734]

Stelmachowicz PG, Jesteadt W, Gorga MP, Mott J. Speech perception ability and psychophysical tuning curves in hearing-impaired listeners. Journal of the Acoustical Society of America. 1985; 77(2):620–627. [PubMed: 3973233]

Stone MA, Glasberg BR, Moore BC. Simplified measurement of auditory filter shapes using the notched-noise method. British Journal of Audiology. 1992; 26(6):329–334. [PubMed: 1292815]

Studebaker G. A “rationalized” arcsine transform. Journal of Speech and Hearing Research. 1985; 28:455–462. [PubMed: 4046587]

Syrdal AK, Gopal HS. A perceptual model of vowel recognition based on the auditory representation of American English vowels. Journal of the Acoustical Society of America. 1986; 79(4):1086–1100. [PubMed: 3700864]

Turner CW, Henn CC. The relation between vowel recognition and measures of frequency resolution. Journal of Speech and Hearing Research. 1989; 32:49–58. [PubMed: 2704201]

Turner CW, Holte LA. Discrimination of spectral-peak amplitude by normal and hearing-impaired subjects. Journal of the Acoustical Society of America. 1987; 81:445–451. [PubMed: 3558961]

Walden TC, Walden BE. Predicting success with hearing aids in everyday living. Journal of the American Academy of Audiology. 2004; 15(5):342–52. [PubMed: 15506496]

Wiley TL, Cruickshanks KJ, Nondahl DM, Tweed TS, Klein R, Klein BE. Tympanometric measures in older adults. Journal of the American Academy of Audiology. 1996; 7(4):260–268. [PubMed: 8827920]

Woods WS, Van Tasell DJ, Rickert ME, Trine TD. SII and fit-to target analysis of compression system performance as a function of number of compression channels. International Journal of Audiology. 2006; 45(11):630–644. [PubMed: 17118906]

Xu L, Pfingst BE. Spectral and temporal cues for speech recognition: implications for auditory prostheses. Hearing Research. 2008; 242(1–2):132–40. [PubMed: 18249077]

Xu L, Thompson CS, Pfingst BE. Relative contributions of spectral and temporal cues for phoneme recognition. Journal of the Acoustical Society of America. 2005 May; 117(5):3255–67. [PubMed: 15957791]

Yund EW, Buckles KM. Multichannel compression hearing aids: Effect of number of channels on speech discrimination in noise. Journal of the Acoustical Society of America. 1995; 97:1206–1223. [PubMed: 7876443]

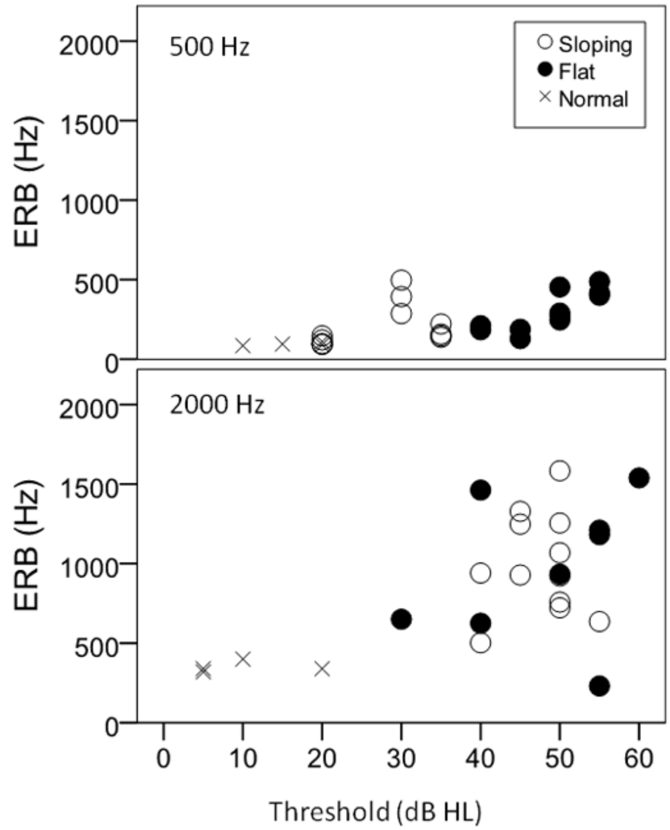

Figure 1.
Auditory filter width (expressed as equivalent rectangular bandwidth [ERB], in Hz) as a function of pure-tone threshold (expressed in dB HL). Results are plotted for 500 Hz in the top panel and for 2000 Hz in the lower panel, for each of the groups. Mean ERB for the listeners with normal hearing was 95 Hz (SD=10.8 Hz) at 500 Hz and 350 Hz (SD=34.6 Hz) at 2000 Hz.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

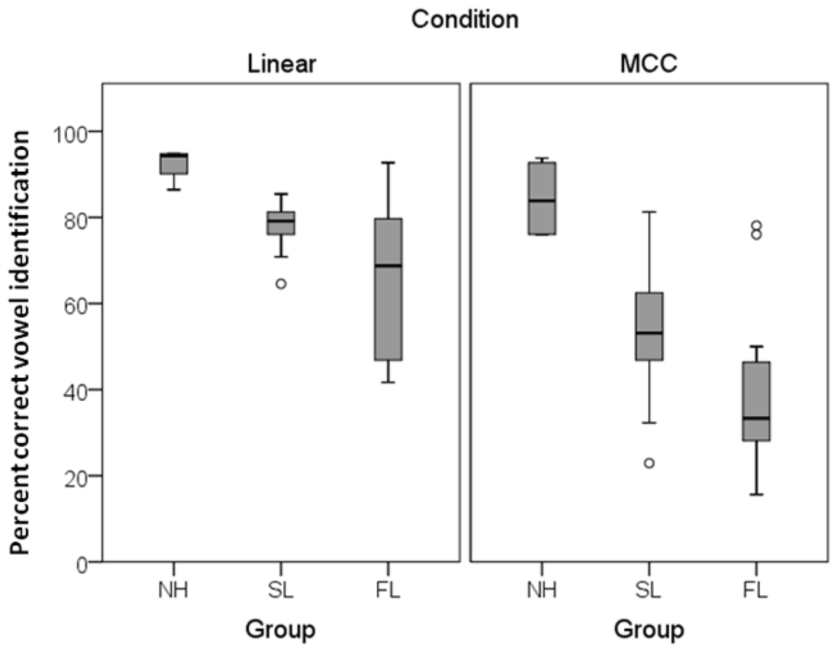

Figure 2.
Vowel identification (in percent correct) for normal hearing (NH), sloping (SL), and flat (FL) listeners. The shaded box represents the 75th to 25th percentile range, and the line in the middle of the box represents the 50th percentile. The whiskers indicate the highest and lowest values that are not outliers. Outliers that fall between 1.5 and 3 times the box length are plotted as open circles.

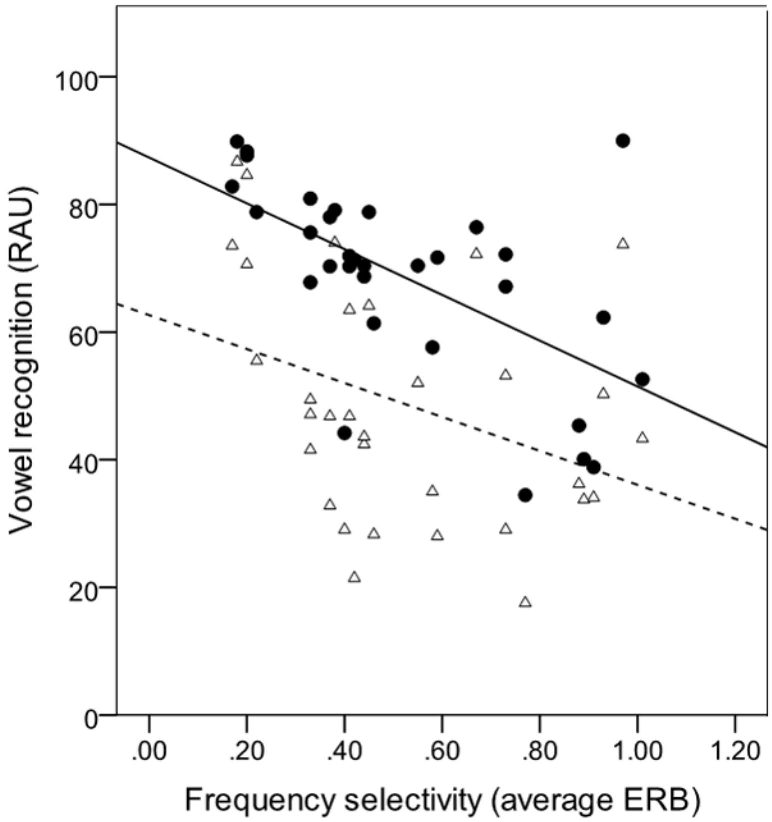

Figure 3.
Relationship between auditory filter and performance. Values on the abscissa are average of normalized ERB at .5 and 2 kHz, where normalized values are calculated as ERB (in Hz) divided by tone frequency (in Hz). Values on the ordinate are vowel identification scores expressed as rationalized arcsine units. Linear data are plotted as filled circles and MCC data as open triangles. The best-fit lines illustrate the trend for linear (solid line) and MCC (dashed line).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

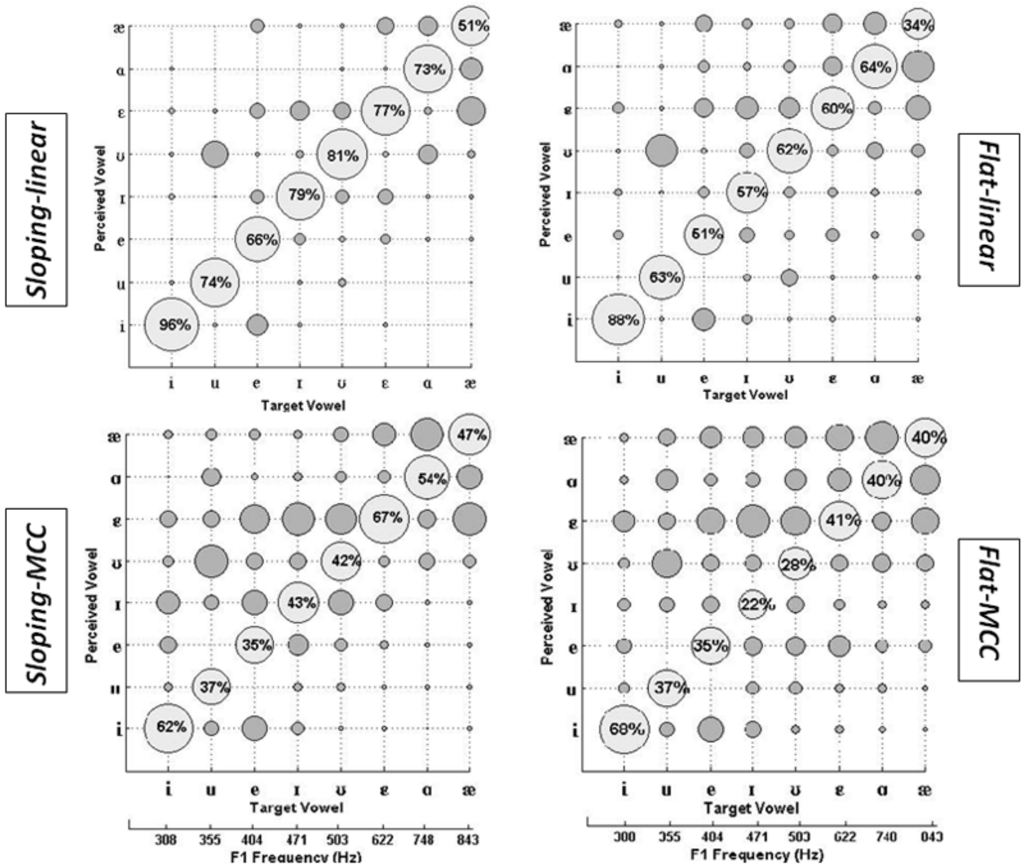

Figure 4.

Vowel error patterns. In each panel, presented vowels, ordered by F1 frequency, are shown on the abscissa. Listener responses are shown on the ordinate. The size of each circle represents the number of responses for a presented/responded combination. Responses on the diagonal are correct, and are noted by percent correct values.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Audiometric thresholds in dB HL (mean and standard error) for the three listener groups.

Frequency (kHz)	**.25**	**.5**	**1**	**2**	**3**	**4**	**6**	**8**
Flat (mean)	44.7	48.0	47.7	48.0	50.7	53.3	50.7	60.7
Flat (standard error)	*2.9*	*1.9*	*2.4*	*2.2*	*3.3*	*3.0*	*2.6*	*4.6*
Sloping (mean)	23.1	27.3	33.8	48.1	50.4	55.8	58.1	63.1
Sloping (standard error)	*2.9*	*1.8*	*1.7*	*1.3*	*1.7*	*2.2*	*2.9*	*4.1*
Normal (mean)	7.5	16.3	15.0	10.0	12.5	12.5	12.5	13.8
Normal (standard error)	*1.4*	*2.4*	*2.0*	*3.5*	*1.4*	*2.5*	*3.2*	*2.4*

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Frequency characteristics of notched noise for auditory filter measurement. The outer edges of the masking noise were fixed at 0.8 x f_c (i.e., 100 and 900 Hz for the 500 Hz tone, and 400 and 3600 Hz for the 2000 Hz tone).

Lower notch edge		Tone	Upper notch edge		Notch width
Hz	**Δ f/f_c**	**f_c (Hz)**	**Δ f/f_c**	**Hz**	**Hz**
500	0.0	500	0.0	500	0
450	0.1	500	0.1	550	100
400	0.2	500	0,2	600	200
300	0.4	500	0.4	700	400
400	0.2	500	0.4	700	300
300	0.4	500	0.2	600	300
2000	0.0	2000	0.0	2000	0
1800	0.1	2000	0.1	2200	400
1600	0.2	2000	0,2	2400	800
1200	0.4	2000	0.4	2800	1600
1600	0.2	2000	0.4	2800	1200
1200	0.4	2000	0.2	2400	1200

Table 3

Mean first and second formant frequencies of male and female vowels).

	Females		Males	
	F1	**F2**	**F1**	**F2**
/i/	327	2991	289	2346
/e/	405	2791	403	2217
/I/	502	2357	441	1942
/ε/	687	2160	557	1791
/æ/	983	1884	703	1622
/ɑ/	817	1259	679	1232
/ʊ/	542	1699	465	1444
/u/	385	1450	324	1223

Table 4

Cutoff frequencies for amplification processing. CF=center frequency.

Channels	Channel	Cutoff Frequency (Hz)	CF (Hz)
Linear	1	141–5623	2882
MCC	1	141 – 178	160
	2	178 – 224	200
	3	224 – 282	250
	4	282 – 355	315
	5	355 – 447	400
	6	447 – 562	500
	7	562 – 708	630
	8	708 – 891	800
	9	891 – 1122	1000
	10	1122 – 1413	1250
	11	1413 – 1778	1600
	12	1778 – 2239	2000
	13	2239 – 2818	2500
	14	2818 – 3548	3150
	15	3548 – 4467	4000
	16	4467 – 5623	5000

Table 5

Bivariate correlation (Pearson r value) for hearing threshold (average of .5 and 2 kHz), relative ERB (average of .5 and 2 kHz), and vowel recognition score.

	Hearing threshold	**Relative ERB**	**Linear vowels**	**MCC vowels**
Hearing threshold		0.67**	−0.65**	−0.64**
Relative ERB	0.67**		−0.60**	−0.36*

*correlation was significant at $p<.05$

**correlation was significant at $p<.01$