Proceedings of Meetings on Acoustics

Volume 26 http://acousticalsociety.org/

171st Meeting of the Acoustical Society of America

Salt Lake City, Utah
23-27 May 2016

Speech Communication: Paper 2aSC14

Spectral contrast effects in vowel categorization by listeners with sensorineural hearing loss

Christian E. Stilp
Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY; christian.stilp@louisville.edu

Joshua M. Alexander
Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN; alexan14@purdue.edu

The auditory system is highly sensitive to changes in acoustic input. This is especially true for sounds with relatively stable (reliable) spectral properties across time. Over a limited range, changes to the spectrum (*e.g.*, spectral peak location upon introduction of a new sound) are perceptually enhanced in proportion to the property's long-term reliability, producing spectral contrast effects (SCEs). For example, a neutral vowel between /ɪ/-/ɛ/ is more likely to be labelled /ɛ/ (high-F_1) when preceded by sounds with a reliable low-frequency peak in the F_1 range for /ɪ/, and *vice versa*. Yet, it is unknown how SCEs affect speech perception by hearing-impaired (HI) listeners because research has only examined normal-hearing (NH) listeners. Here, listeners with mild-to-moderate HI identified target vowels varying from /ɪ/-/ɛ/ that followed a precursor sentence. Reliability of precursor spectral peaks was manipulated using low-F_1 or high-F_1 bandpass filters with +5 to +20 dB gain. SCE magnitude was proportional to precursor filter gain (like NH listeners), and surprisingly, to the amount of low-frequency hearing loss. Thus, mechanisms responsible for SCEs are not dependent on healthy hearing, and are magnified by auditory filter broadening associated with sensorineural hearing loss. Implications of these findings for speech perception will be discussed.

Published by the Acoustical Society of America

INTRODUCTION

The auditory system is highly sensitive to changes in acoustic input, including changes that play an important role in speech perception. One classic demonstration comes from Ladefoged and Broadbent (1957). They varied the first formant frequency (F_1) of a precursor sentence that preceded a test word whose vowel varied perceptually from [bɪt] to [bɛt] and acoustically in F_1 frequency (low and high, respectively). Perception of the vowel was contrastive to the long-term spectrum of the precursor sentence: an emphasis on lower-F_1 frequencies (more /ɪ/-like) in the precursor sentence made the following target vowel sound more like higher-F_1 /ɛ/, and *vice versa*. That is, spectral differences in the F_1 frequency region between the precursor sentence and the target vowel were perceptually magnified. A hallmark characteristic of this phenomenon, known as a spectral contrast effect (SCE), is that categorical judgments of the target sound are biased away from the prominent spectral properties of the sounds that preceded it. These effects have been observed repeatedly for a wide range of spectral properties in speech (Watkins, 1991; Holt, 2005; Sjerps *et al.*, 2011; Stilp *et al.*, 2015) and even in nonspeech sounds (Stilp *et al.*, 2010).

Recently, Stilp and colleagues (2015) revealed that SCE magnitude was a graded response that varied as a function of the long-term spectrum of the precursor. Contrary to prior treatment of SCEs as being simply present or absent, Stilp and colleagues found that SCE magnitude increased with progressively larger spectral peaks in the precursor sentence (creating by increasing filter gain to amplify these key frequency regions). Additionally, for a fixed level of filter gain, wider filters (300 Hz bandwith) produced larger SCEs compared to narrower filters (100 Hz bandwith). By scaling along the bandwidth and/or the magnitude of preceding spectral peaks, SCEs might exert considerable influence on everyday speech perception.

Despite the apparent widespread influence of SCEs on speech perception, to date they have only been considered for listeners with normal-hearing thresholds. One might assume that this is because spectral processing in impaired hearing is insufficient to elicit these context effects in speech perception. For example, listeners with sensorineural hearing loss (SNHL) often have poor frequency resolution due to broadened auditory filter tuning (Pick *et al.*, 1977; Festen & Plomp, 1983; Moore & Glasberg, 1986) and to dysfunction of peripheral mechanisms that produce suppression (Wightman *et al.*, 1977; Sidwell & Summerfield, 1985; Moore & Glasberg, 1986). Abnormal suppression and/or broadened auditory filters make spectral peak detection and vowel recognition more difficult, and also reduce the enhancement of spectral changes (Bacon & Brandt, 1982; Viemeister & Bacon, 1982; Summerfield *et al.*, 1987; Leek *et al.*, 1987; Thibodeau, 1991).

There are multiple reasons to expect that listeners with SNHL demonstrate SCEs in speech perception. First, while sensory encoding differs across normal hearing and SNHL, the rest of the auditory system is similarly predicated on enhancing changes in acoustic input. Enhancement of spectral changes occurs both peripherally (at the cochlea) and centrally. Physiological correlates of psychophysical enhancement (*e.g.*, Viemeister & Bacon, 1982) have been reported at the auditory nerve (Palmer *et al.*, 1995), cochlear nucleus (Scutt & Palmer, 1998), and inferior colliculus (Nelson & Young, 2010). Dichotic stimulus presentation, where preceding acoustic context is presented to the opposite ear of the target sound, still produces SCEs (Watkins, 1991; Holt & Lotto, 2002) and enhancement effects (Erviti *et al.*, 2011; Carcagno *et al.*, 2012). Second, while listeners with SNHL in Thibodeau (1991) did not uniformly show enhancement effects for forward masking of a 2-kHz tone, they did show small but significant enhancement effects for vowel stimuli (one-tailed *t*-tests against chance on data in

Table III of Thibodeau [1991] both $p < .025$). Spectral changes in complex stimuli were enhanced in frequency regions with SNHL, albeit to a lesser degree than for NH listeners. This is consistent with suppression being reduced but not eliminated in moderate hearing loss (Sidwell & Summerfield, 1985; Moore & Glasberg, 1986).

The results of Stilp *et al.* (2015) offer a more specific prediction for how SCEs influence speech categorization by listeners with SNHL. For example, broadening the bandwidth of reliable spectral peaks in the precursor sentence increased SCE magnitude for normal-hearing listeners (Stilp *et al.*, 2015). Broadened auditory filters in hearing-impaired listeners are predicted to have a similar effect, producing larger SCEs than are observed for normal-hearing listeners. The present experiments test this prediction by presenting the same stimuli to normal-hearing listeners and listeners with SNHL.

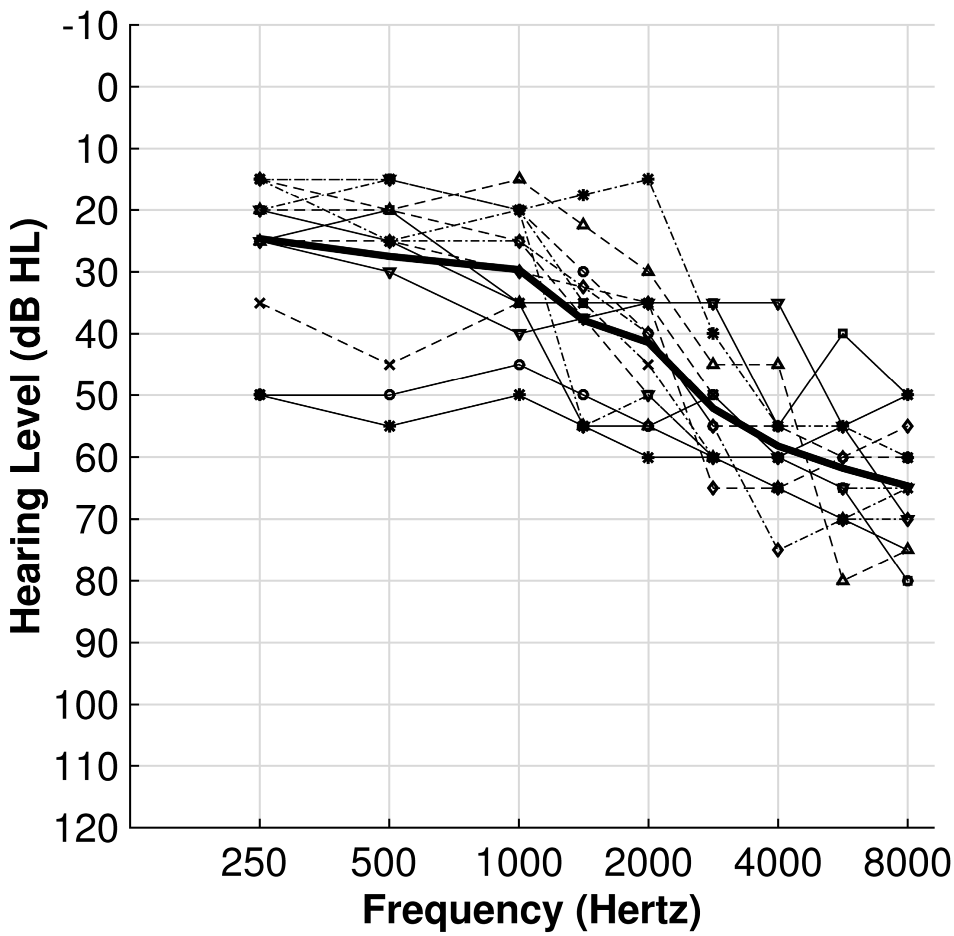

Figure 1: Audiometric thresholds for the test ears of the listeners with SNHL. Each thin line represents a different listener, and the thick black line represents the group mean. The four audiograms with 250- and 500-Hz thresholds at or below the mean form the mild to moderate low-frequency hearing loss subgroup; the remaining audiograms form the near-normal low-frequency hearing subgroup.

METHODS

1. Listeners

A. Hearing-Impaired

Fourteen native-English speaking listeners (ages 51-87) with mild to moderate SNHL from the West Lafayette, Indiana community participated in the experiment. Tympanometry and bone conduction thresholds in each test ear were consistent with normal middle ear function (*i.e.*, compliance peak height $\geq$ 0.3 mmho with peak pressure between -150 and +50 daPa, and most air-bone gaps $\leq$ 10 dB).

B. Normal-Hearing

Twenty-five undergraduate students from the University of Louisville who reported no known hearing loss participated as a comparison group. These listeners were included to replicate and extend findings from Stilp *et al.* (2015), where two different listener groups completed four levels of filter gain.

2. Stimuli

Stimuli are the same as those tested in Stilp *et al.* (2015) and Assgari and Stilp (2015). Natural tokens of /ɪ/ ("ih") and /ɛ/ ("eh") were digitally edited to vary in F_1, forming a ten-step series of target vowels each 246 ms in duration. The precursor was a recording of the first author saying "Please say what this vowel is". The precursor duration was 2174 ms, and measures of long-term average energy in low F_1 (100-400 Hz) and high F_1 (550-850 Hz) regions were within 1 dB of each other. The precursor sentence and all target vowels were set to equal root-mean-square (RMS) amplitudes. Spectral peaks were added to the precursor using FIR bandpass filters in MATLAB with 1200 coefficients. Low F_1 or high F_1 regions were amplified in the precursor sentence by +5, +10, +15, or +20 dB. Precursor sentences and target vowels were then concatenated with a 50-ms silent inter-stimulus interval.

3. Procedure

Listeners sat in sound-treated booths at a computer that controlled stimulus presentation via custom MATLAB scripts. Stimuli were presented monaurally (to the test ear represented in Figure 1 for listeners with SNHL) or diotically (normal-hearing listeners) via circumaural headphones[1]. For normal-hearing listeners, vowel targets were presented at a mean level of 70 dB SPL, and precursor sentences were presented at mean levels that depended on the gain of the bandpass filter: 84.17 (+20 dB filtering condition), 79.40 (+15), 74.88 (+10) and 71.08 (+5) dB SPL.

For the hearing-impaired listeners, the precursors and vowel targets were amplified offline by a hearing aid simulator in MATLAB (see Alexander & Masterson, 2015) that provided linear gain to scale the output to levels prescribed by the Desired Sensation Level *m*(I/O) algorithm v5.0a (National Centre for Audiology, University of Western Ontario) for a 60-dB SPL presentation level. Frequency shaping was accomplished using eight channels whose center and crossover frequencies were based on the recommendations of the DSL algorithm: 0.315, 0.5, 0.8, 1.25, 2.0, 3.15, and 5.0 kHz. The highest-frequency channel (5 kHz to Nyquist, 11.25 kHz) was assigned 0 dB of gain to avoid saturating the output of the headphone. The mean frequency-gain function is shown in Figure 2. The frequency-gain function for all listeners was monotonically increasing up through 2-4 kHz. Because the same frequency shaping was applied to all precursors and targets throughout the entire sequence of testing, it not expected to have influenced target vowel identification or SCEs.

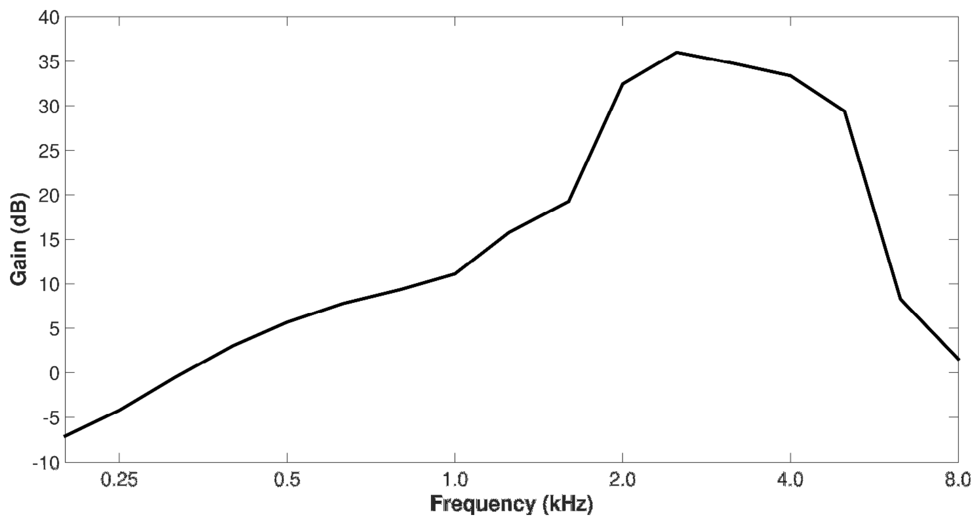

Figure 2: Mean gain at each frequency provided by the hearing aid simulator for the hearing-impaired listeners. For all listeners, gain monotonically increased through 2-4 kHz. As shown in the figure, for low frequencies where hearing is normal or near normal, the Desired Sensation Level algorithm prescribes slightly negative gain to help minimize upward spread of masking. Gain was linear, hence constant, across the presentation of the precursor and target and across all trials.

On each trial, listeners heard a precursor sentence followed by a target vowel. Listeners clicked the mouse or used the touchscreen to report whether the target vowel sounded more like “ih” (as in ‘bit’) or “eh” (as in ‘bet’). For the normal-hearing listeners, trials were arranged into four blocks, with each block testing one level of filter gain (+5, +10, +15, +20 dB). Each block consisted of 160 trials (10 vowels x 2 levels of F_1 filtering x 8 repetitions), which were self-paced. Blocks were completed in random orders, and listeners were allowed to take breaks between blocks as needed. The procedure was modified for the hearing-impaired listeners in anticipation that they would experience initial difficulty with the task. For the hearing-impaired listeners, trials were arranged in 10 blocks, with each block consisting of 20 trials (10 vowels x 2 levels of F_1 filtering). The additional blocks helped to provide them with practice and randomizing the filter gains with each block helped to distribute any effects of practice over the duration of the experiment. These differences are not expected to influence the results. These procedures were approved by the University of Louisville Institutional Review Board (normal-hearing listeners) and the Purdue University Institutional Review Board (hearing-impaired listeners).

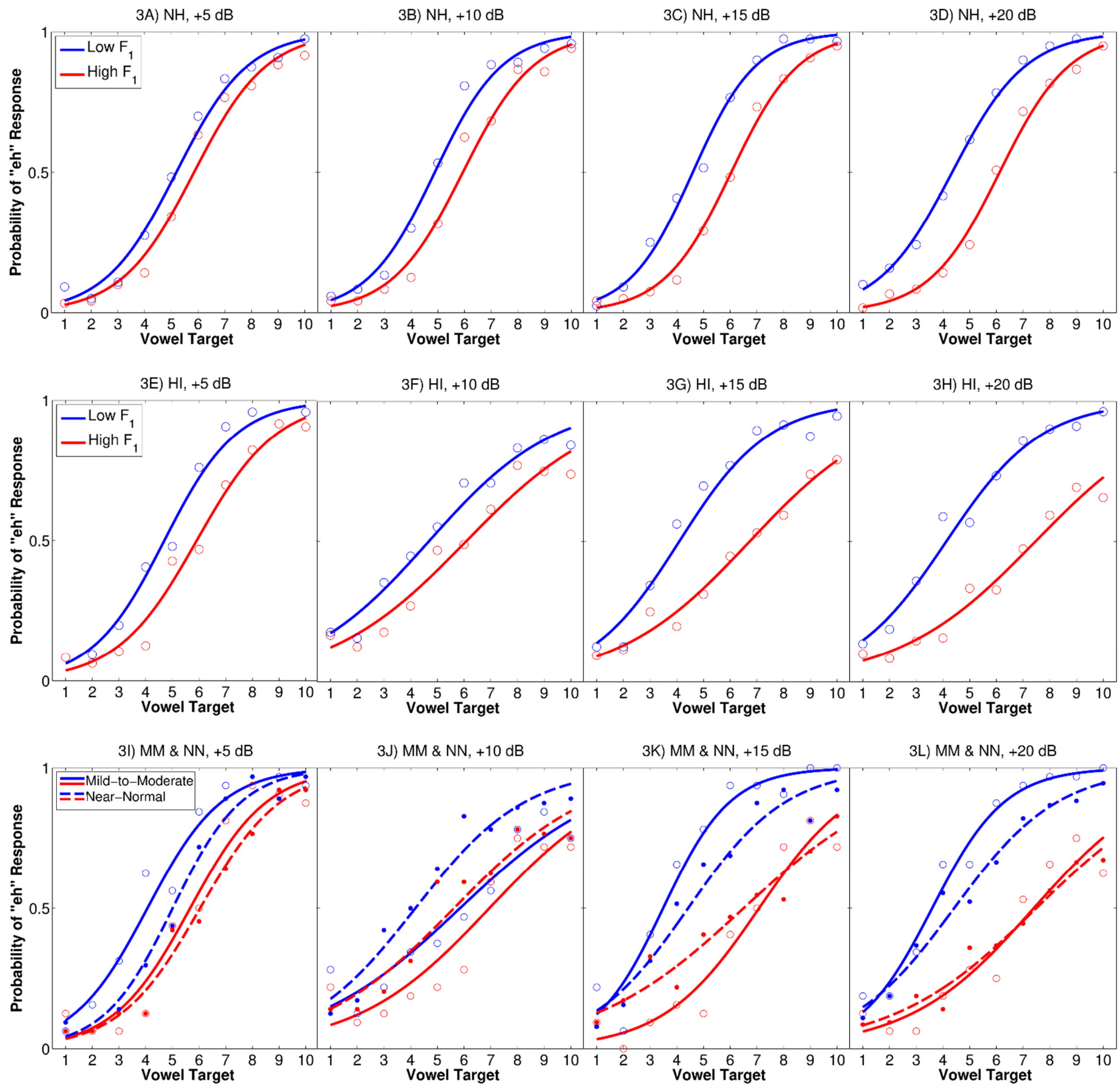

Figure 3: Mean probabilities of "eh" responses are presented as a function of vowel target. Blue and red lines indicate responses following precursor sentences filtered to emphasize low-F_1 and high-F_1 frequency regions, respectively. From left to right, columns show data for filters with +5, +10, +15, and +20 dB of gain. Logistic regressions are fit to mean group data for illustration purposes. The first row (3A-3D) shows mean responses for normal-hearing listeners. The second row (3E-3H) shows mean responses for all hearing-impaired listeners. The third row (3I-3L) shows mean responses for two subgroups of hearing-impaired listeners: those with near-normal low-frequency hearing (dashed lines with solid symbols) and those with mild-to-moderate low-frequency hearing loss (solid lines with open symbols).

RESULTS

Results from five of the normal-hearing listeners were excluded from analysis due to an inability to consistently identify unambiguous vowel stimuli (endpoints of the stimulus series), making interpretation of vowel category boundary shifts due to SCEs problematic. Likewise, results from two hearing-impaired listeners were excluded for similar reasons. This resulted in 20 normal-hearing listeners and 12 hearing-impaired listeners in the final analysis.

For both listener groups, the first two repetitions of each stimulus were treated as practice trials and were not included in statistical analyses. For each listener, for each level of filter gain, logistic regression was used to fit to the vowel identification data associated with each precursor (low- vs. high-F_1 filter peak). For each function, the 50% point was calculated from the regression equation to identify the stimulus that elicited “eh” and “ih” responses with equal probability. These points were often interpolated between members of the vowel series (*e.g.*, stimulus number 4.75 in the series which was numbered discretely from 1 to 10). Following Stilp *et al.* (2015), SCEs were calculated as the difference in the 50% points across the regression fits for each precursor. This metric provided an index for how much vowel perception was influenced by the spectral emphasis in each precursor.

Logistic regression yielded extremely poor fits for two data sets from hearing-impaired listeners (+10 dB condition for one listener; +15 dB condition for another listener). Calculating SCEs using the above method on these ill-fitting logistic functions produced shifts that exceeded 10 stimulus steps, which is greater than the maximum range of the stimuli presented. Examination of each listener’s responses showed no clear evidence of an SCE, as they were approximately equally likely to respond “eh” whether a precursor sentence with a low-F_1 emphasis or a high-F_1 emphasis was presented. Therefore, SCEs for these conditions were set to 0. These listeners exhibited clear SCEs in other stimulus conditions, making the reason for these aberrant results unclear.

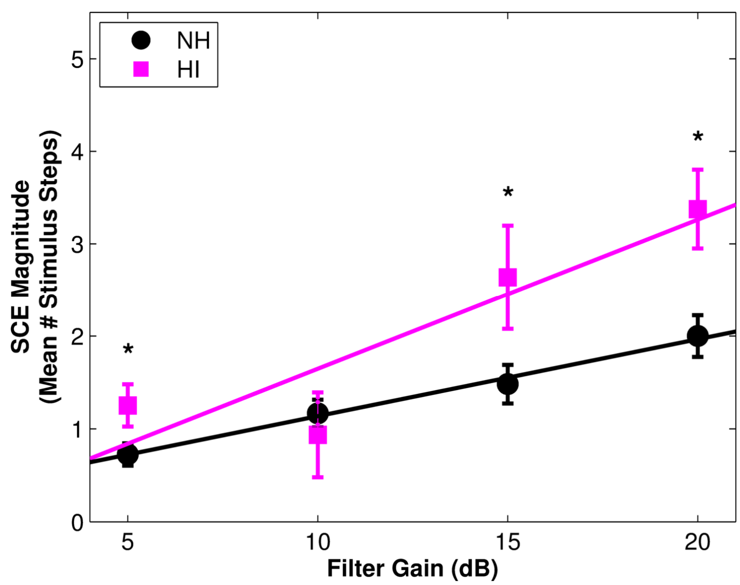

Figure 4: Mean SCE magnitude as a function of filter gain. Results for normal-hearing listeners are shown with black circles; results for hearing-impaired listeners are shown with magenta squares. Solid lines depict linear regression fits for each listener group. Asterisks indicate statistically significant differences across listener groups at that level of filter gain. Error bars indicate standard error of the mean.

Mean SCE magnitude is illustrated as a function of filter gain in Figure 4. These results were analyzed in a 2-by-4 mixed-design ANOVA in SPSS, with listener group as a between-subjects variable (normal hearing, hearing-impaired) and filter gain as a within-subjects variable (+5, +10, +15, +20 dB). Mauchly's test indicated a violation of sphericity, therefore adjusted degrees of freedom and *p* values using the Greenhouse-Geiser correction are reported. The effect of listener group was statistically significant ($F_{1,30} = 7.12$, $p < .025$, $\eta_p^2 = 0.19$), revealing that listeners with SNHL (mean SCE = 2.05 stimulus steps) exhibited larger SCEs than the NH listeners (mean SCE = 1.34 steps). SCE magnitude also varied as a function of filter gain ($F_{2.42,72.50} = 21.28$, $p < .001$, $\eta_p^2 = 0.42$). From Figures 3 and 4, it is readily apparent that larger filter gains produced greater SCEs, consistent with Stilp *et al.* (2015).

The interaction between listener group and filter gain was also statistically significant ($F_{2.42,72.50} = 4.08$, $p < .05$, $\eta_p^2 = 0.12$). This interaction was followed up by an independent-samples *t*-test at each level of filter gain. Except for the +10 dB filter gain condition ($t_{30} = -0.58$, $p = 0.57$), hearing-impaired listeners exhibited significantly larger SCEs than the normal-hearing listeners (+5 dB condition: $t_{30} = 2.25$, $p < 0.05$; +15 dB: $t_{30} = 2.29$, $p < 0.05$; +20 dB: $t_{30} = 3.14$, $p < 0.01$). This pattern of differences is supported by the linear increase in SCE magnitude as a function of filter gain for each listener group. The linearity of this relationship for each group was at or approached statistical significance (normal-hearing: $r = 0.99$, $p < .01$; hearing-impaired: $r = 0.90$, $p = .10$).

While listeners in the hearing-impaired group all displayed impaired high-frequency hearing (Figure 1), this might not be the best predictor of performance for a task where experimental manipulations and target speech cues were at much lower frequencies (< 1000 Hz). Upon examination of the audiometric data, the 12 hearing-impaired listeners formed two subgroups: eight listeners with near-normal low-frequency hearing ($\leq$ 25 dB HL at 250 and 500 Hz, $\leq$ 35 dB HL at 1000 Hz; mean age = 70.63 years) and four listeners with mild-to-moderate low-frequency hearing loss (25-55 dB HL at 250 and 500 Hz, 35-55 dB HL at 1000 Hz; mean age = 66.75 years). Results from Figure 4 are replotted in Figure 5 to highlight the differences in the mean SCEs for these hearing-impaired subgroups.

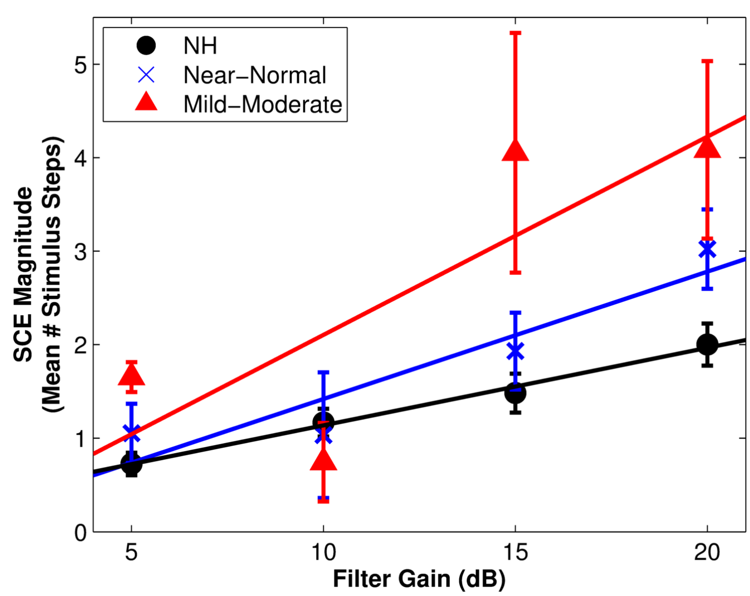

Figure 5: Mean SCE magnitude as a function of filter gain. Results for normal-hearing listeners are shown with black circles; results for hearing-impaired listeners with near-normal low-frequency hearing are shown with blue x's; results for hearing-impaired listeners with mild-to-moderate low-frequency hearing loss are shown with red triangles. Solid lines depict linear regression fits for each listener group. Error bars indicate standard error of the mean.

Results were analyzed in a 3-by-4 mixed-design ANOVA in SPSS, with listener group as a between-subjects variable (normal hearing, near-normal low-frequency hearing-impaired group, mild-to-moderate low-frequency hearing-impaired group) and filter gain as a within-subjects variable. Mauchly's test indicated a violation of sphericity, so adjusted degrees of freedom and p values using the Greenhouse-Geiser correction are reported. The effect of listener group was again statistically significant ($F_{2,29} = 6.09$, $p < .01$, $\eta_p^2 = 0.30$). Independent-samples t-tests indicated that listeners with mild-to-moderate low-frequency hearing loss produced larger SCEs than normal-hearing listeners ($t_{22} = 3.49$, $p < .005$), and other pairwise comparisons were in the predicted directions but not statistically significant (among hearing-impaired groups: $t_{10} = 1.55$, $p = .15$; near-normal low-frequency hearing vs. normal-hearing group: $t_{26} = 1.69$, $p = .10$). SCE magnitude again varied as a function of filter gain ($F_{2.35,68.11} = 24.06$, $p < .001$, $\eta_p^2 = 0.45$), and the interaction between listener group and filter gain was again significant ($F_{4.70,68.11} = 3.68$, $p < .01$, $\eta_p^2 = 0.20$).

Of principal interest are potential differences between the hearing-impaired subgroups. These groups are both small and highly variable, but exploratory analyses can prove informative. Independent-samples t-tests revealed slightly but not significantly larger SCEs for the mild-to-moderate group in the +5 dB condition ($t_{10} = 1.28$, $p = .23$) and no difference between groups in the +10 dB condition ($t_{10} = -0.29$, $p = .78$). SCEs trended toward being significantly larger for the mild-to-moderate group at +15 dB ($t_{10} = 2.03$, $p = .07$). Finally, SCEs were again slightly but not significantly larger for mild-to-moderate group in the +20 dB condition ($t_{10} = 1.20$, $p = .26$). While it is premature to make definitive conclusions with these limited samples, the consistent directionality of SCE differences in the +5, +15, and +20 dB conditions offer a promising pattern of results that warrants future investigation with sufficient samples sizes in both groups.

SCE magnitude was comparable between the normal-hearing group and near-normal group for +5, +10, and +15 dB filter gains (independent-samples t-tests all $t_{26} < 1.20$, $p > .24$), but the near-normal group exhibited significantly larger SCEs at +20 dB filter gain ($t_{26} = 2.29$, $p < .05$). Conversely, the mild-to-moderate group exhibited significantly larger SCEs than did normal-hearing group for +5 dB ($t_{22} = 3.29$, $p < .01$), +15 dB ($t_{22} = 3.66$, $p < .01$) and +20 dB conditions ($t_{22} = 3.25$, $p < .01$). Results did not significantly differ across these groups for the +10 dB condition ($t_{22} = 1.12$, $p = .28$).

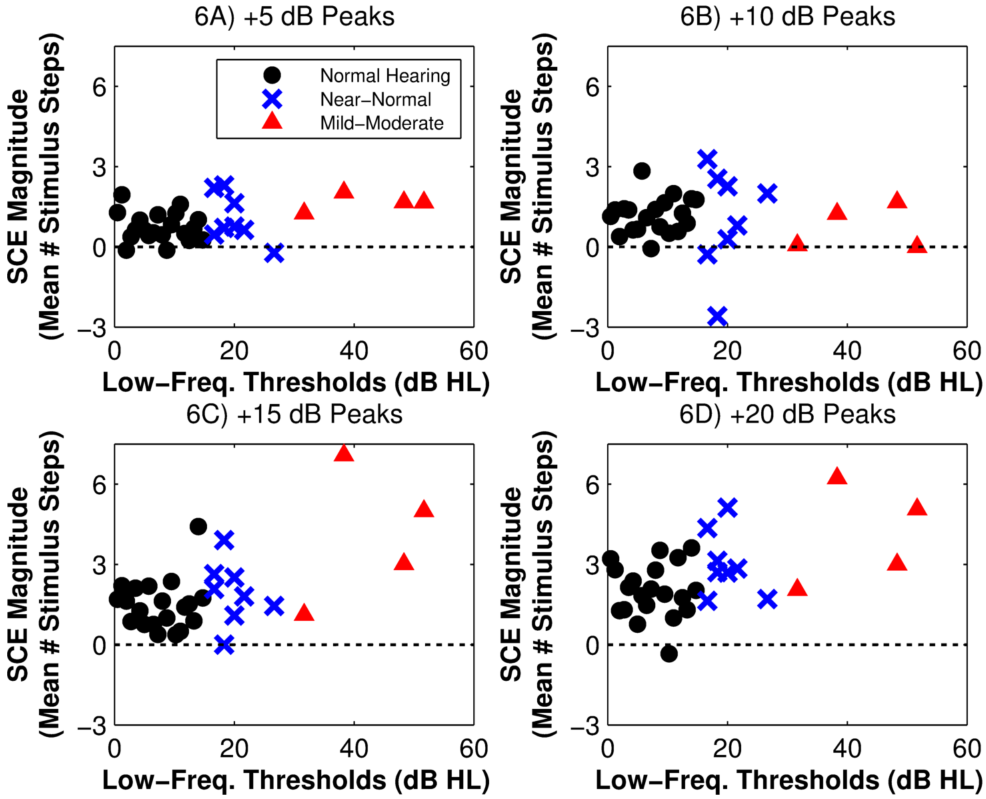

Figure 6: Scatterplots showing SCE magnitude (measured in number of stimulus steps separating response functions) as a function of low-frequency threshold. The normal-hearing group is shown with black circles, the near-normal group with blue x's, and the mild-to-moderate group with red triangles. For both hearing-impaired listener groups, low-frequency hearing loss was computed as the mean of audiometric thresholds at 250, 500, and 1000 Hz, measured in dB HL. Low-frequency hearing was not measured in listeners who reported normal hearing, so their results arbitrarily span 0-15 dB HL for illustration purposes.

Figure 6 shows the variation of low-frequency hearing and SCE magnitude for listeners with SNHL. Low-frequency hearing was calculated as the average audiometric thresholds at 250, 500, and 1000 Hz, thereby spanning the low- and high-F_1 regions manipulated in the present experiment. While listeners with mild-to-moderate low-frequency hearing loss generally exhibited slightly larger SCEs than listeners with near-normal low-frequency hearing (Figure 4), there was no consistent relationship between the extent of low-frequency hearing loss and SCE magnitude across all listeners with SNHL (+5 dB condition: $r = 0.25$, $p = .44$; +10 dB condition: $r = –0.03$, $p = .92$; +15 dB condition: $r = 0.54$, $p = 0.07$; +20 dB condition: $r = 0.35$, $p = .26$).

DISCUSSION

Demonstrations of SCEs across highly variable approaches and stimuli support their widespread influence on speech perception for normal-hearing listeners (see Stilp *et al.*, 2015 for review). The present results reveal for the first time that this influence extends to listeners with SNHL as well. Beyond merely observing SCEs, several important characteristics of hearing-impaired listeners' results are noteworthy. First, SCE magnitude followed the same pattern as established for normal-hearing listeners in Stilp *et al.* (2015): progressively larger spectral peaks in the precursor sentence elicited larger SCEs. Second, SCEs were significantly larger for hearing-impaired listeners than those observed for normal-hearing listeners in most conditions (+5, +15, and +20 dB filter gains). Third, SCE magnitude increased more quickly for hearing-impaired listeners as spectral peak heights in the precursor sentence increased (Figure 4). Hearing-impaired listeners' SCEs in the +10 dB condition were smaller than expected; reasons for this result are unclear. Nevertheless, spectral context effects in speech perception are not exclusive to normal peripheral encoding, and are perhaps qualitatively different (larger overall, increase more rapidly) in listeners with SNHL compared to normal-hearing listeners.

Separating hearing-impaired listeners into two subgroups based on their hearing thresholds near the F_1 frequencies produced illuminating results. Listeners with mild to moderately impaired low-frequency hearing tended to exhibit larger SCEs than listeners with near-normal low-frequency hearing. The small sample sizes of these subgroups reduced the power available for statistical comparisons, but results are nonetheless suggestive of differential processing of contextual spectral information based on degree of hearing impairment.

It is important to clarify the apparent inverse relationship between processing spectral peaks and their effect on speech perception for hearing-impaired listeners. It is well-established that thresholds for detecting spectral peaks are poorer for hearing-impaired listeners than normal-hearing listeners (*e.g.*, Leek *et al.*, 1987). In such experiments, spectral peaks are considered to be *intrinsic* cues to vowel identity as they directly contribute to the acoustics of the vowel target (*e.g.*, Ainsworth, 1975). In the present experiments, reliable spectral peaks in the preceding sentence are considered to be *extrinsic* cues to vowel identity as they occur before the vowel target. These extrinsic, suprathreshold spectral peaks exert larger influences on speech perception (*i.e.*, SCEs) for hearing-impaired listeners than for normal-hearing listeners. The nature of the spectral peaks (intrinsic versus extrinsic cues to vowel identity) as well as their presentation levels (near-threshold versus suprathreshold) warrant close consideration when characterizing their contribution to speech perception for hearing-impaired listeners.

There are a few possible mechanisms that may cause SCEs in regions of hearing loss to be larger-than-normal. First, presentation levels in regions of hearing loss were presented at higher SPLs. It has been demonstrated that auditory filtering broadens as intensity levels increase, even for normal-hearing listeners. This phenomenon would cause spectral peaks to cover a wider excitation region on the cochlea. Stilp *et al.* (2015) showed that SCEs in normal-hearing listeners are larger when induced by spectral peaks that cover a wider bandwidth. Therefore, if broadened auditory filters, whether due to SNHL or to overall presentation level, function to spread the peak energy along the basilar membrane in the same way as wider acoustic filters, then this could explain the differences between the two hearing-impaired subgroups. Another possible mechanism is the steeper growth of loudness that accompanies SNHL. That is, as filter gain is increased, the perceived loudness would grow at a greater rate when the filter is centered in a region with SNHL. This could explain why the SCE magnitude was greater for the hearing-impaired listeners compared to the normal-hearing listeners at higher filter gains (+15 and +20 dB), contributing to steeper regression slopes in Figures 4 and 5.

Given that SCE magnitude tended to be greater in regions of SNHL, it is reasonable to question how this could be pathological and detrimental to speech perception. Spectral context effects have been proposed to be a means of normalizing at least some of the extreme acoustic variability in speech (Wang & Oxenham, 2016). Without this normalization, speech perception may be more difficult and less accurate due to inappropriate processing of acoustic detail. The present results illustrate the opposite case, where context effects exert a larger influence on speech perception than that observed for normal-hearing listeners. This situation also poses a challenge to accurate speech recognition. With overly large SCEs, category boundaries between speech sounds are shifted by larger amounts, increasing the frequency with which listeners give a single response to speech stimuli. For example, across a wide range of the vowel series, hearing-impaired listeners with mild-to-moderate low-frequency hearing loss were more likely to identify the target vowel as "eh" following the low-F_1 precursor than were the near-normal hearing-impaired listeners (Figures 3K-3L), and likewise for responding "ih" following the high-F_1-emphasized precursor (Figure 3K). On a token level, the "eh" endpoint (stimulus 10) was identified consistently by normal-hearing listeners in both F_1 conditions (Figure 3D) and by hearing-impaired listeners in the low-F_1 condition (Figure 3H). However, the increased influence of SCEs led to only about 65% "eh" responses to this stimulus in the high-F_1 condition (Figures 3H, 3L). Overly large SCEs made typically unambiguous vowels more ambiguous and were thus miscategorized more often.

Results also have great importance for digital signal processing approaches in hearing aids and cochlear implants. Perception of a given speech sound is the product of acoustic information in that sound, the sound that comes immediately before (*e.g.*, short-term SCEs, coarticulation), and sounds from the past few seconds (*e.g.*, long-term SCEs, normalizing to talker, accent, dialect, room acoustics, *etc.*). Speech sound recognition by hearing-impaired listeners is influenced by both short-term and long-term properties of the listening context (Alexander & Kluender, 2009). This argues strongly against the exclusive use of short time constants in hearing aid filtering (*e.g.*, Van Dijkhuizen *et al.*, 1987, 1989), as this would neglect the influence of earlier spectral information shown here to influence vowel perception by hearing-impaired listeners.

ENDNOTES

[1] There is no empirical precedent for SCE magnitudes differing across diotic and monotic presentation. Dichotic presentation, where precursor and target sounds are presented to opposite ears, appeared to produce smaller SCEs than diotic presentation (Watkins, 1991), but dichotic presentation was not utilized in the present study. Therefore, no specific predictions are offered for how SCE magnitudes might differ based on diotic (normal-hearing listeners) versus monotic presentation (hearing-impaired listeners).

REFERENCES

Ainsworth, W. (1975). "Intrinsic and extrinsic factors in vowel judgments," in *Auditory Analysis and Perception of Speech*, edited by G. Fant and M. Tatham (Academic, London), pp. 103-113.

Alexander, J. M., & Kluender, K. R. (2009). Spectral tilt change in stop consonant perception by listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research*, *52(3)*, 653-670.

Alexander, J. M., & Masterson, K. M. (2015). Effects of WDRC release time and number of channels on output SNR and speech recognition. *Ear and Hearing*, *36(2)*, e35-e49.

Bacon, S. P., & Brandt, J. F. (1982). Auditory processing of vowels by normal-hearing and hearing-impaired listeners. *Journal of Speech, Language, and Hearing Research*, *25(3),* 339-347.

Carcagno, S., Semal, C., & Demany, L. (2012). Auditory enhancement of increments in spectral amplitude stems from more than one source. *Journal of the Association for Research in Otolaryngology*, *13(5),* 693-702.

Erviti, M., Semal, C., & Demany, L. (2011). Enhancing a tone by shifting its frequency or intensity. *Journal of the Acoustical Society of America*, *129(6),* 3837-3845.

Festen, J.M., & Plomp, R. (1983) Relations between auditory functions in impaired hearing. *Journal of the Acoustical Society of America 73(2)*, 652–662.

Holt, L. L. (2005). Temporally nonadjacent nonlinguistic sounds affect speech categorization. *Psychological Science*, *16(4),* 305-312.

Holt, L. L., & Lotto, A. J. (2002). Behavioral examinations of the level of auditory processing of speech context effects. *Hearing Research*, *167(1),* 156-169.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *The Journal of the Acoustical Society of America*, *29(1)*, 98-104.

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(1)*, 148-154.

Moore, B.C.J., & Glasberg, B.R. (1986) Comparisons of frequency selectivity in simultaneous and forward masking for subjects with unilateral cochlear impairments. *Journal of the Acoustical Society of America, 80(1),* 93–107.

Nelson, P. C., & Young, E. D. (2010). Neural correlates of context-dependent perceptual enhancement in the inferior colliculus. *The Journal of Neuroscience*, *30(19)*, 6577-6587.

Palmer, A. R., Summerfield, Q., & Fantini, D. A. (1995). Responses of auditory-nerve fibers to stimuli producing psychophysical enhancement. *The Journal of the Acoustical Society of America*, *97(3),* 1786-1799.

Pick, G., Evans, E. F., & Wilson, J. P. (1977). "Frequency resolution in patients with hearing loss of cochlear origin", in *Psychophysics and Physiology of Hearing*, eds. Y. Cazals, K. Horner, & L. Demany (Pergamon Press: New York) pp. 273-281.

Scutt, M.J., and Palmer, A.R. (1998). Physiological enhancement in cochlear nucleus using single tone precursors. *Association for Research in Otolaryngology Abstracts, 381(A).*

Sidwell A., & Summerfield Q. (1985). The effect of enhanced spectral contrast on the internal representation of vowel-shaped noise. *Journal of the Acoustical Society of America, 78(2)*, 495–506.

Sjerps, M. J., Mitterer, H., & McQueen, J. M. (2011). Constraints on the processes responsible for the extrinsic normalization of vowels. *Attention, Perception, & Psychophysics*, *73*(4), 1195-1215.

Stilp, C.E., Alexander, J.M., Kiefte, M., & Kluender, K.R. (2010). Auditory color constancy: Calibration to reliable spectral properties across speech and nonspeech contexts and targets. *Attention, Perception & Psychophysics, 72(2)*, 470-480.

Stilp, C. E., Anderson, P. W., & Winn, M. B. (2015). Predicting contrast effects following reliable spectral properties in speech perception. *The Journal of the Acoustical Society of America*, *137*(6), 3466-3476.

Summerfield, Q., Sidwell, A., & Nelson, T. (1987). Auditory enhancement of changes in spectral amplitude. *Journal of the Acoustical Society of America, 81(3),* 700-708.

Thibodeau, L. M. (1991). Performance of hearing-impaired persons on auditory enhancement tasks. *The Journal of the Acoustical Society of America*, *89(6)*, 2843-2850.

Van Dijkhuizen, J. N., Anema, P. C., & Plomp, R. (1987). The effect of varying the slope of the amplitude-frequency response on the masked speech-reception threshold of sentences. *The Journal of the Acoustical Society of America*, *81*(2), 465-469.

Van Dijkhuizen, J. N., Festen, J. M., & Plomp, R. (1989). The effect of varying the amplitude-frequency response on the masked speech-reception threshold of sentences for hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *86*(2), 621-628.

Viemeister, N.F., & Bacon, S.P. (1982). Forward masking by enhanced components in harmonic complexes. *Journal of the Acoustical Society of America, 71(6),* 1502-1507.

Wang, N., & Oxenham, A.J. (2016). Effects of auditory enhancement on the loudness of masker and target components. *Hearing Research*, *333*, 150-156.

Watkins, A. J. (1991). Central, auditory mechanisms of perceptual compensation for spectral envelope distortion. *Journal of the Acoustical Society of America, 90(6),* 2942-2955.

Wightman, F., McKee T., & Kramer M. (1977) “Factors influencing frequency selectivity in normal and hearing impaired listeners” in *Psychophysics and Physiology of Hearing*, eds. E.F. Evans and J.P. Wilson (Adacemic Press: London) pp. 295–310.