Published in final edited form as:

J Speech Lang Hear Res. 2010 October ; 53(5): 1087–1095. doi:10.1044/1092-4388(2010/09-0185).

Influence of Broad Auditory Tuning on Across-Frequency Integration of Speech Patterns

Eric W. Healy and

The Ohio State University, Columbus, and University of South Carolina, Columbia

Kimberly A. Carson

University of South Carolina

Abstract

Purpose—The purpose of the present study was to assess whether diminished tolerance for disruptions to across-frequency timing in listeners with hearing impairment can be attributed to broad auditory tuning.

Method—In 2 experiments in which random assignment was used, sentences were represented as 3 noise bands centered at 530, 1500, and 4243 Hz, which were amplitude modulated by 3 corresponding narrow speech bands. To isolate broad tuning from other influences of hearing impairment, listeners with normal hearing (45 in Experiment 1 and 30 in Experiment 2) were presented with these vocoder stimuli, having carrier band filter slopes of 12, 24, and 192 dB/octave. These speech patterns were presented in synchrony and with between-band asynchronies up to 40 ms.

Results—Mean intelligibility scores were reduced in conditions of severe, but not moderate, simulated broadening. Although scores fell as asynchrony increased, the steeper drop in performance characteristic of listeners with hearing impairment tested previously was not observed in conditions of simulated broadening.

Conclusions—The intolerance for small across-frequency asynchronies observed previously does not appear attributable to broad tuning. Instead, the present data suggest that the across-frequency processing mechanism in at least some listeners with hearing impairment might be less robust to this type of degradation.

Keywords

hearing impairment; speech perception; across frequency

The abilities and limitations of individuals having a sensorineural hearing loss are of profound importance and, accordingly, have formed the topic of a considerable amount of study. The known limitations include a loss of audibility, the broadening of frequency tuning (cf. Moore, 2007), the loss of compressive nonlinearity (cf. Bacon, Fay, & Popper, 2004; Horwitz, Ahlstrom, & Dubno, 2007), and an especially steep growth of loudness (Fowler, 1936; Steinberg & Gardner, 1937). In contrast to these more established limitations, the ability to process speech patterns across various frequency separations is more poorly understood in listeners with hearing impairment (HI).

A number of recent studies have suggested that listeners with HI may possess some limitation in their ability to combine the information provided by temporal speech envelopes that are separated in frequency. These studies have often used vocoder processing, in which speech is divided into frequency bands, the temporal envelope of each is used to modulate a corresponding noise or tone carrier, and the modulated carriers are presented to listeners (Dudley, 1939; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). It has generally been found that performance of listeners with normal hearing (NH) and HI is similar when recognition is based on a single temporal envelope (e.g., Turner, Souza, & Forget, 1995). This suggests that the ability to process speech envelope information is intact in the impaired auditory system. However, as the number of spectral channels increases, so that recognition involves integration across various spectral loci, listeners with HI generally perform more poorly than those with NH (Başkent, 2006 [in noise]; Grant, Tufts, & Greenberg, 2007; Healy & Bacon, 2002; Healy, Kannabiran, & Bacon, 2005; Souza & Boike, 2006; Turner, Chi, & Flock, 1999). The exception to this pattern of results is provided by Hall, Buss, and Grose (2008b), who used sentences filtered to band-widths that provided 15%–25% recognition at each of two center frequencies. When the bands were presented simultaneously, scores were similar across listeners with NH and those with mild to moderate HI.¹

Healy and Bacon (2002) used stimuli that differed somewhat from most of the studies listed above. Sentence materials were filtered to 1/3-octave bands centered at 750 and 3000 Hz, which were then used to modulate pure tones at matching frequencies. The use of this widely separated narrow band pair was reasoned to have the following advantages: (a) because intelligibility resulting from the individual narrow bands was essentially zero, any intelligibility observed must have resulted from a combination or integration of the band pair, (b) audibility for the listeners with HI could be tailored and carefully controlled for each listener, and (c) interactions between the spectral bands resulting from broad auditory tuning should have been reduced.

Two primary results were obtained. The first involved reduced recognition by the listeners with HI in a baseline condition in which the two bands were presented in synchrony. The second result, which is of particular importance for the present work, involves the tolerance that listeners with NH and HI had to a disruption in the across-frequency timing information provided by the band pair. When an asynchrony was introduced between the bands, such that one led the other by some tens of milliseconds, the listeners with HI displayed a steeper than normal drop in performance. In fact, a subgroup of listeners with HI who produced reasonably high scores when the bands were presented synchronously demonstrated a pronounced drop in performance in response to a slight asynchrony (12.5ms) that yielded essentially no reduction in the listeners with NH. In a subsequent study, Healy et al. (2005) examined this effect further. The reductions in performance as a function of asynchrony were compared with a measure of acoustic signal disruption caused by asynchrony. This measure involved the reduction in correlation between the acoustic temporal patterns as asynchrony increased. The authors found that the rate of performance reduction observed for the listeners with NH closely matched the amount of signal disruption. However, the performance of the listeners with HI fell at a rate that was disproportionate to the signal disruption. It was concluded that the listeners with HI tested by Healy and Bacon (2002) demonstrated deficient across-frequency processing, as reflected by a reduced ability to process intact (synchronous) across-frequency speech information and a lack of tolerance for a slight disruption in across-frequency cues produced by asynchrony.

---

¹Hall, Buss, and Grose (2008a) found that listeners with HI performed better when speech cues were distributed over time versus when they were clustered in time (asynchronous vs. synchronous “band AM”). However, as the authors pointed out, this does not constitute a conflict with the findings or conclusions of Healy and Bacon (2002) because of the large differences in the paradigms used in the two studies.

Although the conditions used by Healy and Bacon (2002) were designed to provide some control for limitations known to accompany HI, it remains important to examine the specific influence of these known limitations on across-frequency processing. An example is provided by Souza and Boike (2006), who found no evidence that listener age caused the across-frequency processing difficulties observed by Healy and Bacon (2002) in listeners with HI.

Another known limitation that could potentially influence across-frequency processing involves broad auditory tuning. Due to a loss of the active mechanism provided by the outer hair cells, listeners having a cochlear hearing loss of roughly 40 dB or greater often possess broader than normal frequency tuning (cf. Moore, 2007). Studies in which the influence of broadened tuning on speech reception is examined have used spectral smearing using the overlap-add method (Allen & Rabiner, 1977) or a reduction in filter slope angles to produce overlap between adjacent spectral bands. In summary, this work has shown that speech reception in quiet is somewhat resistant to moderate amounts of smearing or overlap. However, when the communication setting is made more challenging through the use of background noise, or when smearing becomes severe, recognition can be reduced (Baer & Moore, 1993; Boothroyd, Mulhearn, Gong, & Ostroff, 1996; Celmer & Bienvenue, 1987; Fu & Nogaki, 2005; Healy & Warren, 2003; Litvak, Spahr, Saoji, & Fridman, 2007; Shannon, Zeng, & Wygonski, 1998; ter Keurs, Festen, & Plomp, 1992, 1993; Villchur, 1977).

However, it is not known to what extent spectral broadening can affect the tolerance for disruption to across-frequency timing produced by across-channel asynchrony. There is some reason to believe that broadening might affect this aspect of across-frequency processing. Healy and Bacon (2007) found that asynchrony produced a larger intelligibility reduction for band pairs that were closely spaced in frequency relative to band pairs that were widely spaced. The effect was not attributable to mixing of the asynchronous speech patterns in the region of spectral overlap, indicating that overlap between the bands was not responsible. However, it was hypothesized that across frequency timing information might be more readily available to the auditory system when that information is derived from spectral regions having smaller frequency separation. Thus, it remains possible that broadening of speech patterns might be similar to reducing center frequency spacing in its tendency to reduce tolerance for across-frequency timing disruption.

The present study was designed to test the specific hypothesis that broad auditory tuning can account for the steeper than normal reduction in performance displayed by listeners with HI as a function of between band asynchrony in Healy and Bacon (2002). Although across-channel asynchrony may not commonly occur in natural listening, it may serve as a tool to examine the robustness of the across-frequency processing mechanism. In a pair of experiments, listeners having NH were used for their sharp uniform tuning and to eliminate other potentially confounding aspects of HI. A simulation of broad tuning involving modification of filter slopes in a noise vocoder was used to provide control over the amount of broadening in each condition.

Experiment 1 most closely resembled the conditions used by Healy and Bacon (2002) and others. The filter slopes used to create speech bands (sometimes referred to as *analysis bands*) were fixed across conditions, but the temporal envelopes derived from these regions were presented to spectral regions having various amounts of broadening. Experiment 2 more closely resembled speech processing by the impaired auditory system, albeit with a reduced number of channels. In this experiment, analysis and presentation-band broadening was matched so that speech envelopes were presented to the spectral region from which they were derived.

Experiment 1

Method

Subjects—Forty-five university students received course credit for participation. All were native speakers of English and ranged in age from 18 to 39 years (mean age of all subjects tested = 22 years). Each subject had audiometric thresholds of 20 dB HL or better bilaterally at octave frequencies from 250 to 4000 Hz (ANSI, 1996), and none had previous exposure to the sentence materials used in these experiments. The study protocol was approved by the Institutional Review Board at the University of South Carolina.

Stimuli—The 100 everyday American speech sentences from the Central Institute for the Deaf (CID; Davis & Silverman, 1978; Silverman & Hirsh, 1955) were used. Ten practice sentences were chosen from the Speech Perception in Noise (SPIN) Test (Kalikow, Stevens, & Elliot, 1977). The sentences were recorded at 16-bit resolution and 22-kHz sampling using a professional speaker having a standard American dialect.

To construct the noise-vocoder stimuli, the speech signal was first filtered to three 1/3-octave bands having center frequencies of 530, 1500, and 4243 Hz (analysis bands). Four cascaded passes through eighth-order Butterworth filters were used, yielding slopes of 192 dB/octave. Temporal information was extracted from each band using half-wave rectification and low-pass filtering (250-Hz cutoff, 100-order fir filter) and was used to amplitude modulate white noise using sample point-by-point multiplication (Hori, House, & Hughes, 1971). Presentation-band filtering was then used to restrict each amplitude-modulated noise, using center frequencies that matched the corresponding speech bands. The parameter of interest was the presentation-band filter slope. Three separate sets of conditions were created in which slopes were 12, 24, or 192 dB/octave (single passes through second- or fourth-order, or four cascaded passes through eighth-order Butterworth filters, respectively). These filter slope conditions were selected to represent the varying degrees of broadening shown in Figure 1.

The level of each sentence in each band was equated to ensure that the peak of the slow response root-mean-square averages were equal. Prior to mixing into three-band stimuli, the frequency-specific delays associated with filtering were assessed empirically by filtering a single sample click. These observed delays were compensated for by applying the following corrections: for the 12-dB/octave stimuli, Bands 2 and 3 (1500 and 4243 Hz) were delayed relative to Band 1 (530 Hz) by 13 and 17 ms, respectively; for the 24-dB/octave stimuli, Bands 2 and 3 were delayed relative to Band 1 by 13 and 16 ms, respectively; and for the 192-dB/octave stimuli, Bands 2 and 3 were delayed relative to Band 1 by 24 and 33 ms, respectively. The bands comprising each presentation-band filter slope condition were then mixed at equal amplitude.

Once the bands were synchronized, as described above, additional conditions were created in which the bands were mixed in a staggered asynchronous fashion. The delays, expressed relative to Band 1, ranged up to 40 ms and are displayed in Table 1. Thus, there were five asynchrony conditions (including 0-0-0 ms) in each of the three filter slope conditions. All manipulation was performed digitally, primarily in MATLAB. A PC and Echo Gina 24 D/A converters were used to present the digital files. The signal was routed through a Mackie 1202VLZ mixer for convenient calibration and monitoring. The analog signal was presented diotically at a slow peak level of 70 dBA over Sennheiser HD 250 Linear II headphones.

The present stimuli differed from those used by Healy and Bacon (2002) in two primary ways. First, pure-tone carriers were replaced with noise carriers because spectral broadening of the carrier is only possible with the latter. Second, an additional frequency band was added, for a total of three, because sentence intelligibility based on narrow-band temporal patterns is

generally lower for noise than for pure-tone carriers (Healy & Warren, 2003), and scores needed to be in an appropriate range above the performance floor.

Procedure—Subjects were randomly assigned to one of the three broadening conditions. Subjects in each group ($n = 15$) heard 17 sentences in each of the five asynchrony conditions. The sentence-to-condition correspondence was balanced to ensure that each sentence was heard in each condition an equal number of times, and the order in which asynchrony conditions were heard was randomized for each subject. The session began with practice consisting of the 10 SPIN sentences heard first broadband, then again in the synchronous condition. This was followed by 15 CID practice sentences in the same condition. Subjects then heard the 85 remaining CID sentences in the various test conditions. Each subject was tested individually in a sound-attenuating booth seated with the experimenter. They were instructed to repeat back as much of each sentence as they could and were encouraged to guess if unsure. Each test sentence was presented only once. The experimenter controlled the presentation of the stimuli and recorded each correctly recalled keyword.

Results and Discussion

Mean intelligibility scores and standard errors as a function of asynchrony value for the three presentation-band broadening conditions are shown in Figure 2. Performance was greatest in the steep and moderate slope conditions (192 and 24 dB/octave) but was substantially poorer in the broadest condition (12 dB/octave). As expected, performance fell as a function of between-band asynchrony. However, most importantly, no indication of a steeper drop in performance with the introduction of a small asynchrony in the broadest condition appears relative to the other conditions.

These observations are supported by statistical examination. A two-factor mixed analysis of variance (ANOVA) on rationalized arcsine-transformed values (RAU; Studebaker, 1985) revealed significant main effects of broadening condition, $F(2, 42) = 22.9, p < .001$, and asynchrony, $F(4, 168) = 135.0, p < .001$, but no interaction between these factors, $F(8, 168) = 1.7, p = .11$. Post hoc analyses (Holm-Sidak, $\alpha = .05$, throughout) of the main effect of broadening revealed that scores in the 192- and 24-dB/octave conditions were equivalent but differed from the 12-dB/octave condition. Post hoc analyses of the main effect of asynchrony indicated that performance at each asynchrony differed. Repeated measures ANOVA conducted separately for each broadening condition indicated a significant main effect of asynchrony in the 192-dB/octave condition, $F(4, 56) = 35.6, p < .001$; the 24-dB/octave condition, $F(4, 56) = 53.1, p < .001$; and the 12-dB/octave condition, $F(4, 56) = 54.3, p < .001$. Post hoc analyses indicated that performance was reduced similarly, at an asynchrony of 10 ms relative to synchrony, in each broadening condition.

One difference between the results of the present study and those of Healy and Bacon (2002) is that significant declines in performance were observed at an asynchrony of 10 ms in the present experiment, whereas no decline was observed previously in listeners with NH at a similar asynchrony. This difference is likely related to differences in stimuli used in the two studies. In accord with Healy and Bacon (2007), a larger influence of between-band asynchrony can be observed with a smaller, relative to a larger, separation in center frequency. Thus, the smaller band separation necessitated by the use of a third spectral band in the present experiment likely caused performance to be reduced at a smaller asynchrony than in the 2002 study. Although performance is reduced as a function of a small asynchrony in the present experiment, importantly, this reduction is the same in all broadening conditions.

The present data suggest that a moderate broadening of speech patterns (24 dB/octave) produces no reduction in performance. However, a more severe broadening of 12 dB/octave was sufficient to produce a significant reduction in intelligibility. It is worth noting that 12 dB/

octave approximates the slopes of auditory filters of individuals having rather severe levels of broad auditory tuning (cf. Moore, 2007, Figure 3.22, p. 81). This result is therefore in accord with other reports in the suggestion that a severe, but not a moderate, broadening/spectral smearing can indeed reduce intelligibility, especially if speech materials have been subjected to some degradation prior to broadening/smearing. However, the present data do not support the hypothesis that broad auditory tuning can account for the steeper than normal reduction in performance displayed by listeners with HI as a function of between-band asynchrony in Healy and Bacon (2002).

Experiment 1 resembled Healy and Bacon (2002) in that analysis filtering, and therefore the amount of speech information, was held constant across broadening conditions. To further investigate a possible relationship between the steep drop in performance as a function of asynchrony and broadened tuning, a second experiment was conducted in which conditions were designed to more closely resemble those found in the impaired auditory system. In Experiment 2, the same filter parameters were used for both analysis and presentation-band filtering. Thus, speech patterns were derived from and presented to the same spectral regions, which again had three levels of broadening.

Experiment 2

Method

Thirty additional listeners were recruited, having the characteristics described in Experiment 1. The speech materials and processing techniques were identical to those of Experiment 1, except as noted. Two new conditions were created for comparison to the 192-dB/octave presentation-band conditions from Experiment 1. In these new conditions, presentation-band filtering was again 12 or 24 dB/octave. However, instead of 192-dB/octave analysis band filtering, analysis filtering parameters matched presentation-band filtering. To correct for the filter delays, Bands 2 and 3 of the 24-dB/octave stimuli were delayed relative to Band 1 by 2 and 3 ms, respectively. The 12-dB/octave stimuli required no correction to retain alignment across frequency bands.

Each sentence in each band was equated in amplitude, and the bands comprising each condition were mixed in synchrony and at the four asynchronies used in Experiment 1. Presentation to the two groups of listeners at 70 dBA was accomplished using the same apparatus used previously. Furthermore, all test procedures, including practice with SPIN sentences followed by 17 CID test sentences in each condition, were identical to those in Experiment 1.

Results and Discussion

Figure 3 shows mean intelligibility in the new conditions (24 and 12 dB/octave), along with scores in the 192-dB/octave condition from Experiment 1 (replotted from Figure 2). As in Experiment 1, scores are lowest in the broadest condition (12 dB/octave). However, scores in this experiment are highest in the conditions having moderate broadening (24 dB/octave). Also in accord with the results of Experiment 1, scores fell as a function of between-band asynchrony. But again, the rate at which scores fell was similar across conditions of broadening.

A two-factor mixed ANOVA performed on the RAU-transformed data revealed significant main effects of broadening condition, $F(2, 42) = 13.4, p < .001$, and asynchrony, $F(4, 168) = 108.9, p < .001$, and a nonsignificant interaction, $F(8, 168) = 1.2, p = .33$. Post hoc analyses of the main effect of broadening condition indicated that scores in all broadening conditions differed, and analyses of the main effect of asynchrony indicated that all scores differed. Repeated measures ANOVA conducted separately for the two new broadening conditions indicated a significant main effect of asynchrony at 24 dB/octave, $F(4, 56) = 43.7, p < .001$, and at 12 dB/octave, $F(4, 56) = 32.2, p < .001$. Post hoc analyses indicated that performance

was reduced similarly, at an asynchrony of 10 ms relative to synchrony, in each broadening condition.

The nonmonotonic relationship between spectral broadening and overall level of performance in the present experiment may be understood in terms of a trade-off between the advantage provided by the additional speech information passed by the shallower filters and the disadvantage associated with broad overlapping spectral patterns. As the slopes of the speech analysis filters are made shallower, larger amounts of speech information are passed, thus increasing scores. However, as broadening increases, the increase in speech information is eventually offset by the severe smearing and near complete spectral overlap of this information.

A similar trade-off may exist in Experiment 1, in which both the 12- and 24-dB/octave conditions exhibit considerable spectral overlap (see Figure 1) yet produce substantially different overall levels of performance. It appears that the increase in stimulation provided by the moderately broad bands (24 dB/octave), even though they possess no additional speech information, is sufficient to offset any potential detriment associated with their spectral overlap.

General Discussion

In Experiment 1, analysis filtering was held constant as presentation-band filtering was broadened so that speech information was constant across broadening conditions. Because these conditions simulated the presentation of narrow frequency band temporal patterns to impaired listeners who likely had various levels of broad auditory tuning, they were more similar to the conditions of Healy and Bacon (2002) than were those of Experiment 2. This experiment therefore provided the critical test of the hypothesis that the steeper than normal drop in performance as a function of asynchrony observed by Healy and Bacon (2002) was attributable to broad tuning. Experiment 2 extended these observations by including conditions that more closely simulated processing by the impaired ear. Because a steeper drop in performance as a function of asynchrony in conditions of simulated broad tuning was not observed in either experiment, the results of the present study indicated that the intolerance for across-channel asynchrony in listeners with HI observed by Healy and Bacon (2002) is likely not attributable to broad auditory tuning.

The listeners with HI tested by Healy and Bacon (2002) fell into two subgroups: those who performed poorly in a baseline condition in which temporal speech patterns appeared in synchrony and those who performed similarly to listeners with NH in synchrony but exhibited an especially steep drop in performance as asynchrony was introduced. It is possible to speculate that the listeners in the former subgroup might have been those with the very broadest auditory tuning, and their performance was reduced in the baseline condition as a result. However, those in the latter subgroup might have had more normal peripheral tuning but suffered from some other limitation that prohibited the tolerance to across-frequency disruptions displayed by the listeners with NH.

The primary manipulations in the present study are spectral. The use of steep narrow-band analysis filtering simulated to some extent the extraction of speech information during normal peripheral filtering. In accord with normal filtering, speech cues were derived from restricted spectral regions. This information was then presented either to matching restricted regions or to regions of simulated broadening. Alternatively, the use of broad analysis filters resulted in speech information averaged across a wide spectral region. In these conditions, spectrally distinct speech information was not extracted, just as it is not extracted in the impaired auditory system.

However, in addition to these more obvious spectral effects, differences between the two experiments exist in the temporal domain. Figure 4 shows temporal speech patterns output by

simulated auditory filters having slopes matching those used in the present study. As can be seen in the upper panel, steep analysis filters output temporal patterns that are limited in fluctuation rate as a result of relatively narrow effective bandwidth. The top two panels show that not only was the spectral range from which speech information was obtained held constant as the presentation band was broadened, but the temporal patterns were also preserved and simply presented to broader spectral regions of excitation. The figure also shows (lower panel) that broad analysis filters produce temporal speech patterns containing abnormally high fluctuation rates. Thus, as the filters were broadened in Experiment 2, speech information was obtained from broad spectral regions usually occupied by several auditory filters (the spectral effect), and the output temporal patterns contained abnormally high fluctuation rates (the temporal effect). Although the higher rate temporal patterns were smoothed to some extent by normal auditory filtering in the present study, the internal representations of the stimuli in Experiments 1 and 2 certainly differed. The observation of similar results across the two experiments strengthens the present conclusions.

Although much is known about the processing abilities and limitations of listeners with HI, across-frequency processing in these listeners remains somewhat less clear. As mentioned earlier, a number of studies in which speech stimuli were used (with the exception of Hall et al., 2008b) have shown reduced performance by these listeners on tasks requiring the processing of temporal speech patterns at different spectral loci. However, these same listeners generally perform quite well on tasks thought to reflect across-frequency processing of *nonspeech* stimuli. *Envelope correlation*, in which listeners identify correlated versus uncorrelated noise envelopes at different spectral loci, is similar across listeners with NH and HI (Hall & Grose, 1993). *Modulation detection interference* (MDI), in which the threshold for detecting amplitude modulation of a carrier is increased when a second modulated carrier is introduced, is essentially normal in individuals with HI (Bacon & Opie, 2002; Grose & Hall, 1994). *Comodulation masking release* (CMR), in which the threshold for a signal in a modulated masker is reduced when a second modulated masker is introduced, can be reduced in listeners with HI, although this effect may reflect broadened auditory tuning (Hall, Davis, Haggard, & Pillsbury, 1988; Hall & Grose, 1989). Finally, across-frequency gap detection, in which the markers preceding and following a silent gap have different frequency content (Grose, Hall, & Buss, 2001), and across-frequency duration discrimination (Grose, Hall, & Buss, 2004), in which the signals in each interval have different frequency content, are both essentially normal in listeners with HI. This discrepancy between speech and nonspeech studies remains unknown and serves to motivate additional investigation of the across-frequency processing abilities and limitations of listeners with HI.

The finding that listeners with HI were less tolerant of small across-frequency asynchronies than were listeners with NH (Healy & Bacon, 2002) suggested that the across-frequency processing mechanism is less robust in these listeners. The present data indicate that this intolerance is not attributable to broad auditory tuning. Instead, these data support the view that a potential deficit in the across-frequency processing mechanism of at least some listeners with HI exists in addition to known limitations.

Acknowledgments

This work was supported in part by National Institutes of Health Grant DC08594. The assistance of Karin Jacobus and Kimberlee Crass is gratefully acknowledged, as are comments from Frederic Apoux on an earlier version of this article.

References

- Allen JB, Rabiner LR. A unified approach to short-time Fourier analysis and synthesis. *Proceedings IEEE* 1977;65:1558–1564.

- American National Standards Institute. Specifications for audiometers (ANSI S3.6-1996). New York, NY: Author;
- Bacon, SP.; Fay, RR.; Popper, AN. Compression: From cochlea to cochlear implants. New York, NY: Springer; 2004.
- Bacon SP, Opie JM. Modulation detection interference in listeners with normal and impaired hearing. *Journal of Speech, Language, and Hearing Research* 2002;45:392–402.
- Baer T, Moore BCJ. Effects of spectral smearing on the intelligibility of sentences in noise. *The Journal of the Acoustical Society of America* 1993;94:1229–1241.
- Başkent D. Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels. *The Journal of the Acoustical Society of America* 2006;120:2908–2925. [PubMed: 17139748]
- Boothroyd A, Mulhearn B, Gong J, Ostroff J. Effects of spectral smearing on phoneme and word recognition. *The Journal of the Acoustical Society of America* 1996;100:1807–1818. [PubMed: 8817914]
- Celmer, RD.; Bienvenue, GR. Critical bands in the perception of speech signals by normal and sensorineural hearing loss listeners. In: Schouten, MEH., editor. *The psychophysics of speech perception*. Dordrecht, the Netherlands: Martinus Nijhoff; 1987. p. 473–480.
- Davis, H.; Silverman, SR. *Hearing and deafness*. 4th ed. New York, NY: Holt, Rinehart, and Winston; 1978.
- Dudley H. The vocoder. *Bell Labs Record* 1939;17:122–126.
- Fowler EP. A method for the early detection of otosclerosis: A study of sounds well above threshold. *Archives of Otolaryngology–Head & Neck Surgery* 1936;24:731–741.
- Fu Q-J, Nogaki G. Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology* 2005;6:19–27. [PubMed: 15735937]
- Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. *Hearing Research* 1990;47:103–138. [PubMed: 2228789]
- Grant KW, Tufts JB, Greenberg S. Integration efficiency for speech perception within and across sensory modalities by normal-hearing and hearing-impaired individuals. *The Journal of the Acoustical Society of America* 2007;121:1164–1176. [PubMed: 17348537]
- Große JH, Hall JW III. Modulation detection interference (MDI) in listeners with cochlear hearing loss. *Journal of Speech and Hearing Research* 1994;37:680–686. [PubMed: 8084198]
- Große JH, Hall JW III, Buss E. Gap duration discrimination in listeners with cochlear hearing loss: Effects of gap and marker duration, frequency separation, and mode of presentation. *Journal of the Association for Research in Otolaryngology* 2001;2:388–398. [PubMed: 11833611]
- Große JH, Hall JW III, Buss E. Duration discrimination in listeners with cochlear hearing loss: Effects of stimulus type and frequency. *Journal of Speech, Language, and Hearing Research* 2004;47:5–12.
- Hall JW III, Buss E, Große JH. The effect of hearing impairment on the identification of speech that is modulated synchronously or asynchronously across frequency. *The Journal of the Acoustical Society of America* 2008a;123:955–962. [PubMed: 18247898]
- Hall JW III, Buss E, Große JH. Spectral integration of speech bands in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America* 2008b;124:1105–1115. [PubMed: 18681600]
- Hall JW III, Davis AC, Haggard MP, Pillsbury HC. Spectro-temporal analysis in normal-hearing and cochlear-impaired listeners. *The Journal of the Acoustical Society of America* 1988;84:1325–1331. [PubMed: 3198867]
- Hall JW III, Große JH. Spectrotemporal analysis and cochlear hearing impairment: Effects of frequency selectivity, temporal resolution, signal frequency, and rate of modulation. *The Journal of the Acoustical Society of America* 1989;85:2550–2562. [PubMed: 2745879]
- Hall JW III, Große JH. Monaural envelope correlation perception in listeners with normal hearing and cochlear impairment. *Journal of Speech and Hearing Research* 1993;36:1306–1314. [PubMed: 8114497]

- Healy EW, Bacon SP. Across-frequency comparison of temporal speech information by listeners with normal and impaired hearing. *Journal of Speech, Language, and Hearing Research* 2002;45:1262–1275.
- Healy EW, Bacon SP. Effect of spectral frequency range and separation on the perception of asynchronous speech. *The Journal of the Acoustical Society of America* 2007;121:1691–1700. [PubMed: 17407905]
- Healy EW, Kannabiran A, Bacon SP. An across-frequency processing deficit in listeners with hearing impairment is supported by acoustic correlation. *Journal of Speech, Language, and Hearing Research* 2005;48:1236–1242.
- Healy EW, Warren RM. The role of contrasting temporal amplitude patterns in the perception of speech. *The Journal of the Acoustical Society of America* 2003;113:1676–1688. [PubMed: 12656400]
- Horii Y, House AS, Hughes GW. A masking noise with speech-envelope characteristics for studying intelligibility. *The Journal of the Acoustical Society of America* 1971;49:1849–1856. [PubMed: 5125732]
- Horwitz AR, Ahlstrom JB, Dubno JR. Speech recognition in noise: Estimating effects of compressive nonlinearities in the basilar-membrane response. *Ear and Hearing* 2007;28:682–693. [PubMed: 17804982]
- Kalikow DN, Stevens KN, Elliot LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America* 1977;61:1337–1351. [PubMed: 881487]
- Litvak LM, Spahr AJ, Saoji AA, Fridman GY. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *The Journal of the Acoustical Society of America* 2007;122:982–991. [PubMed: 17672646]
- Moore, BCJ. *Cochlear hearing loss: Physiological, psychological and technical issues*. 2nd ed. West Sussex, England: Wiley; 2007.
- Shannon RV, Zeng F-G, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. *Science* 1995 October 13;270:303–304. [PubMed: 7569981]
- Shannon RV, Zeng F-G, Wygonski J. Speech recognition with altered spectral distribution of envelope cues. *The Journal of the Acoustical Society of America* 1998;104:2467–2476. [PubMed: 10491708]
- Silverman SR, Hirsh IJ. Problems related to the use of speech in clinical audiometry. *Annals of Otolaryngology, Rhinology, and Laryngology* 1955;64:1234–1244.
- Souza PE, Boike KT. Combining temporal-envelope cues across channels: Effects of age and hearing loss. *Journal of Speech, Language, and Hearing Research* 2006;49:138–149.
- Steinberg JC, Gardner MB. The dependence of hearing impairment on sound intensity. *The Journal of the Acoustical Society of America* 1937;9:11–23.
- Studebaker GA. A “rationalized” arcsine transform. *Journal of Speech and Hearing Research* 1985;28:455–462. [PubMed: 4046587]
- ter Keurs M, Festen JM, Plomp R. Effect of spectral envelope smearing on speech reception. I. *The Journal of the Acoustical Society of America* 1992;91:2872–2880. [PubMed: 1629480]
- ter Keurs M, Festen JM, Plomp R. Effect of spectral envelope smearing on speech reception. II. *The Journal of the Acoustical Society of America* 1993;93:1547–1552. [PubMed: 8473608]
- Turner CW, Chi S-L, Flock S. Limiting spectral resolution in speech for listeners with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research* 1999;42:773–784.
- Turner CW, Souza PE, Forget LN. Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners. *The Journal of the Acoustical Society of America* 1995;97:2568–2576. [PubMed: 7714274]
- Villchur E. Electronic models to stimulate the effect of sensory distortions on speech perception by the deaf. *The Journal of the Acoustical Society of America* 1977;62:665–674. [PubMed: 903508]

Figure 1.
Amplitude spectra of the three noise-band stimuli having presentation-band filter slopes of 192, 24, or 12 dB/octave.

Figure 2.

Group mean intelligibility scores and standard errors for sentences represented by temporal patterns at three spectral loci. In each of three spectral broadening conditions, the speech information was constant. Relative timing between the bands is indicated. Separate groups of 15 subjects each were used for the three broadening conditions.

Figure 3.

Group mean intelligibility scores and standard errors for sentences represented by temporal patterns at three spectral loci in each of three spectral broadening conditions. Analysis band filtering matched presentation-band filtering, and relative timing between the bands is indicated. Separate groups of 15 subjects each were used for the three broadening conditions.

Figure 4.

Waveforms resulting from speech passed through simulated peripheral filters. The single equivalent rectangular band (Glasberg & Moore, 1990) width filter was centered at 1 kHz and had either steep or shallow slopes (192 or 12 dB/octave). Top panel: Steep analysis and presentation-band filtering produced temporal patterns limited in fluctuation rate and characteristic of normal peripheral filtering. Middle panel: The normal temporal pattern resulting from steep analysis filtering was not disrupted by subsequent shallow presentation-band filtering. However, this normal pattern was presented to a broad region of the spectrum. Bottom panel: Shallow analysis and presentation-band filtering produced temporal patterns

containing abnormally high rates, derived from and delivered to abnormally broad spectral regions.

Table 1

Asynchrony values (milliseconds) used for the five asynchrony conditions (in rows).

Band		
1 (530 Hz)	2 (1500 Hz)	3 (4243 Hz)
0	0	0
0	5	10
0	10	20
0	15	30
0	20	40