Correction of the Peripheral Spatio-Temporal Response Pattern:

A Potential New Signal-Processing Strategy

Lu-Feng Shi

Institute for Sensory Research

Department of Communication Sciences and Disorders

Syracuse University

Syracuse, New York 13244

Laurel H. Carney

Institute for Sensory Research

Department of Bioengineering and Neuroscience

Syracuse University

Syracuse, New York 13244

and

Karen A. Doherty

Institute for Sensory Research

Department of Communication Sciences and Disorders

Syracuse University

Syracuse, New York 13244

Running head: Spatiotemporal pattern correction as signal processing

The purpose of this study is to introduce the potential application of a new signal-processing strategy, spatiotemporal pattern correction (SPC), which is based on our knowledge of the level-dependent temporal response properties of auditory-nerve (AN) fibers in normal and impaired ears. SPC manipulates the temporal aspects of different frequency channels of sounds in an attempt to compensate for the loss of nonlinear properties in the impaired ear. Quality judgments and intelligibility measures of speech processed at various SPC strengths were obtained on a group of normal-hearing listeners and listeners with hearing loss. In general, listeners with hearing loss preferred sentences with some level of SPC processing, whereas normal-hearing listeners preferred the quality of the unprocessed sentences. Benefit from SPC on the nonsense syllable test varied greatly across phonemes and listeners. These preliminary findings suggest that SPC, a temporally based algorithm designed to improve the perception of speech for listeners with hearing loss, has potential to be useful to listeners with hearing loss. However, before this strategy can be integrated in hearing aids, a more comprehensive study on the benefit of SPC for listeners with different degrees and configurations of hearing loss is needed.

INTRODUCTION

Spatiotemporal pattern correction (SPC) is a signal-processing strategy based on the nonlinear properties of the cochlea. It is known that normal-hearing listeners have sharp peripheral filters, whereas filters are much broader in listeners with hearing loss (e.g., Florentine, Buus, Scharf, & Zwicker, 1980; Moore, 1985; Turner & Henn, 1989; Nelson, 1991; Leek & Summers, 1993; Dubno & Schaefer, 1995; Moore, Vickers, Plack, & Oxenham, 1999; Oxenham & Bacon, 2003). When peripheral filters change their shape with input level, the phase properties of the filters also change (Fig. 1). In normal-hearing listeners, tuning is sharp for low-level input sounds, and broadens as the input level increases. These dynamic changes in tuning between low- and high-level input sounds may play a role in normal-hearing listeners' loudness perception and frequency selectivity. In listeners with hearing loss, the sharpness of tuning degrades with increases in hearing loss. The tuning in an ear with mild to moderate cochlear impairment for low-level input sounds is broader than in a normal ear. Tuning in an impaired ear at levels near threshold resembles tuning in a normal ear for high-level input sounds (e.g., Florentine et al., 1980; Moore, 1985; Nelson, 1991). The broadening of filters in the impaired ear has been attributed to damage in outer hair cell (OHC) function (Dallos & Harris, 1978) and has been shown to decrease the recognition of vowels (e.g., Turner & Henn, 1989; Richie, Kewley-Port, & Coughlin, 2003) and/or consonants (e.g., Preminger & Wiley, 1985; Dubno & Dirks, 1989; Dubno & Shaefer, 1995; Turner, Chi, & Flock, 1999).

INSERT FIGURE 1 HERE.

The bandwidth of a filter also affects the phase properties that are related to the latency of the filter's response, or to its group delay. The relationship between group delay and phase properties is illustrated in Fig. 2. The duration of the build-up of a cochlear filter's response depends upon how sharply tuned the filter is. Broad filters (i.e., for high SPLs in a normal ear and for both low and high SPLs in an impaired ear) have short build-up times, whereas sharp filters (i.e., for low SPLs in a normal ear) have a long build-up time. The build-up time is proportional to the group delay. In the normal ear, the actual group delay constantly fluctuates between the low- and high-SPL group-delay values. In the impaired ear, the group delay varies much less across SPLs.

INSERT FIGURE 2 HERE.

In listeners with hearing loss, the lack of the dynamic change in phase over input level could explain some of their poor differentiation of subtle contrasts embedded in speech. The most common approach used in the hearing-aid industry to compensate for the reduction in the nonlinear properties of the impaired ear is wide-dynamic-range-compression (WDRC). This level-based strategy, however, does not compensate for the loss of nonlinearity due to reduced phase delays between low- and high-level input sounds.

WDRC has been widely accepted as an efficient and effective signal-processing strategy. It is a gain-based strategy in that it provides more gain for low input levels than for high input levels. It is designed to improve loudness perception and to ensure that the long-term variation of speech sounds is maintained within a range most comfortable to the listener (e.g., Boothroyd, Springer, Smith, & Schulman, 1988). Because of the nature of compression, the range of output intensity is narrow in WDRC instruments regardless of the input level. As a result, there is a reduction in spectral peak-to-valley contrasts in speech (e.g., Lippman, Braida, & Durlach, 1981; Van Tasell, 1993; Stelmachowicz, Kopun, Mace, Lewis, & Nittrouer, 1995; Hickson, Thyer, & Bates, 1999; Souza & Bishop, 1999; Hedrick & Rice, 2000; Souza & Kitch, 2001). This loss of contrast in dynamic cues changes the relative amplitude between vowels and consonants and reduces speech recognition for listeners with hearing loss (e.g., Summerfield, 1987; Stone & Moore, 1992; Van Tasell, 1993; Souza & Kitch, 2001), especially for high-level speech inputs (Studebaker & Sherbecoe, 1995) and for high WDRC compression ratios (Lippman et al., 1981; Van Tasell & Trine, 1996; Souza & Turner, 1999). This problem is conceivably most prominent in listeners with severe to profound loss, because they require high gain and/or strong compression.

SPC, on the other hand, introduces different delays across frequency channels in the input sound in an attempt to "correct" the abnormal spatiotemporal response pattern without changing the magnitude spectrum of the sound. The delay is introduced so that responses for low- versus high-level input sounds in an impaired cochlea will be more like those in a normal cochlea. Although both WDRC and SPC attempt to correct for the loss of nonlinearities in the impaired cochlea, the approach of each is very different. WDRC is gain-based, whereas SPC is based on temporal information. Thus, there is also the potential that the two strategies may provide greater benefit when combined.

In the current paper we describe a new physiologically-based signal-processing strategy, SPC, and evaluate how listeners with normal hearing and with hearing loss perceive the quality and intelligibility of SPC-processed speech. This paper is the first investigation to assess the feasibility of a signal-processing strategy based on nonlinear temporal properties. Benefit in listeners' performance due to SPC would suggest that the new signal-processing strategy has the potential to be implemented into future hearing-aid technology.

METHOD

Subjects

A total of 18 listeners (6 normal-hearing and 12 listeners with sensorineural hearing loss) participated in the current study. Normal-hearing listeners (2 male, 4 female) were 20 to 57 years of age and had hearing thresholds less than 20 dB HL at the octave frequencies between 250 and 4000 Hz (ANSI, 1989). Of the 12 listeners with hearing loss (5 male, 7 female), 24 to 83 years of age, 10 had a mild to moderate sloping sensorineural hearing loss and 2 had a mild to severe mixed hearing loss, which was consistent with their case history, middle-ear immittance measures, and air- and bone-conduction results. See Table 1 for individual listener's hearing thresholds.

INSERT TABLE 1 HERE.

Three normal-hearing listeners and ten listeners with hearing loss participated in Experiment 1. Data from one listener with hearing loss was excluded from Experiment 1 because the listener could not perform the task. In Experiment 2 four normal-hearing listeners and five listeners with hearing loss participated. One normal-hearing listener and three listeners with hearing loss were participants in both experiments.

SPC Signal Processing

The SPC system is schematically illustrated in Fig. 3. The dynamic time delays for each frequency channel were computed in the following manner:

INSERT FIGURE 3 HERE.

The dynamic temporal properties of healthy auditory-nerve (AN) fibers associated with a given frequency channel were computed using a nonlinear AN model with compression (Heinz, Zhang, Bruce, & Carney, 2001). The dynamic parameters of the AN filters specify both the magnitude and phase properties of the filters as a function of time (Fig. 1). The slope of the phase vs. frequency function for a filter is proportional to its group delay, or cochlear filter build-up time. The group delay is a measure of the overall delay of a signal that passes through the filter due to the tuning of the filter. Group delay is related to bandwidth; thus, this delay is a fundamental temporal property that changes with sound level in the normal ear. This calculation specifies the dynamic temporal properties of the normal ear, which serve as a reference for SPC.

The strength of the SPC applied depended on the assumed loss of nonlinearity in the impaired ear. Sounds were corrected for different degrees of hearing loss; for simplicity, hearing loss was characterized in terms of the percentage of remaining nonlinear function of the impaired ear. The group delay for an impaired filter is always smaller than that of a healthy filter, because broad filters have shorter build-up times. Thus, the appropriate correction is always an inserted delay. The temporal correction was simply a fraction of the normal group delay. This dynamic temporal correction was computed for every time point during the stimulus and for each frequency channel.

The SPC system consists of two signal-processing paths (Fig. 3):

In one path, the time-varying temporal delay for each frequency channel is computed. The use of gammatone filters in the AN model results in very simple group-delay calculations, because the slope of the gammatone filter’s phase-versus-frequency function is simply proportional to the gain of the filter. Gammatone filters provide an excellent description of AN fiber tuning at low and mid frequencies (Carney and Yin, 1988).

In the other path, the correction (a time- and frequency-dependent delay) is inserted between the two stages of an analysis-synthesis filterbank (Hohmann, 2002). The analysis-synthesis filterbank is critical for obtaining high quality signals when combining sounds across different frequency channels. Because each frequency channel is purposefully distorted by the time-varying temporal delays, the final signal is not a reconstruction of the input, but one with spatiotemporal manipulations that are designed to correct the response of the impaired ear. Thus, only listeners with hearing loss can assess the benefit of this system. However, normal-hearing listeners were included in this study to guard against possible artifactual measures of benefit due to unintended aspects of the complex signal manipulations.

Stimuli were pre-processed with several different SPC strengths. Each SPC strength was proportional to a given reduction in the loss of cochlear nonlinearity. For example, to correct for an ear with 80% of normal cochlear nonlinearities, the SPC process introduced 20% of the normal time-varying delay to compensate for the impairment. Relating the percent of normal cochlear nonlinearity directly to a specific degree of hearing loss is difficult to estimate at this stage of the study. Therefore, listeners were tested for a range of SPC strengths to determine a “best” strength. SPC strength was based on 100/ (% assumed normal cochlear nonlinearity); thus the SPC strength for an impaired ear with 80% of normal cochlear nonlinear function is 100/80 or 1.25. Note that in this study the same SPC strength was used to compute corrections for all frequency channels, and each listener was tested with the same range of SPC strengths, regardless of their degree of cochlear impairment.

For the results presented here, the SPC system's analysis filterbank had 2 filters per equivalent rectangular bandwidth (ERB) from 100 to 5000 Hz. The SPC scheme was applied to the filters with center frequencies from 100 to 2000 Hz (i.e., 36 filters). All stimuli were processed using MatLab and C with a 33-kHz sampling rate. All speech stimuli were presented at the input to the SPC system at 65 dB SPL (i.e., conversational speech level); processed sounds were presented to subjects at different SPLs (see below).

Procedures

Listeners were seated in a double-walled sound booth and tested in the sound field. All speech stimuli were presented through a Dell PC and Tucker-Davis Technologies (TDT) DSP board. A programmable attenuator (TDT PA4) and Crown D-75A amplifier were used to control the stimulus level.

In Experiment 1, a two-alternative forced choice (2-AFC) paradigm was employed. Four sentences from the Hearing-in-Noise Test (HINT), spoken by a male speaker in quiet (Nilsson, Soli, & Sullivan, 1994), served as the stimuli. Two versions of the same sentence processed at two SPC strengths with no more than a 0.15 strength difference were presented to a listener on each trial. Listeners were instructed to compare the stimuli in the two intervals and verbally report which one they preferred. They also described the basis for their preference judgments. Before the start of Experiment 1, listeners were given 18 practice trials to familiarize them with the task. Each listener was randomly presented a total of 126 – 432 trials of sentence pairs at 40 dB SL re: speech recognition threshold (SRT). The level was adjusted when listeners reported it was not comfortable. However, the adjusted presentation levels (60 - 85 dB SPL) were always above the listener's SRT and below their uncomfortable loudness level (UCL). To assess if listeners' preference changed with presentation level, two listeners with hearing loss were also presented the stimuli at 45 dB SPL. No differences were observed across presentation levels and therefore data was collapsed across levels for analysis.

In Experiment 2, listeners were randomly presented with one of sixteen vowel-consonant (VC) syllables spoken by a female speaker, a subset of the Nonsense Syllable Test (NST) (Levitt & Resnick, 1978), at five different SPC strengths (1.0, 1.075, 1.15, 1.225, and 1.3), where an SPC of 1.0 indicates that the stimulus was unprocessed. In Experiment 1, correction strengths greater than 1.3 were perceived as highly distorted by both normal-hearing listeners and listeners with hearing loss. The VC stimuli were the vowel /i/ coupled with one of the following sixteen English consonants: /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /θ /, /ð/, /s/, /z/, /ʃ/, /ʒ/, /m/, and /n/.

Listeners participated in a total of four runs (i.e., 1280 trials) in Experiment 2. A single run consisted of 320 trials (16 consonants × 5 correction strengths × 4 repetitions). The total of 1280 trials was collected in one 2 - 3.5 hour listening session. The VCs were presented at 66.2 dB SPL for normal-hearing listeners and varied from 81.8 - 97.8 dB SPL for listeners with hearing loss. Presentation levels never exceeded a listener's UCL.

Listeners were instructed to press one of sixteen buttons on a response box that corresponded to the VC they heard and verbally rate the clarity of the signal on a ten-point scale. This scale was based on the Judgment of Sound Quality (JSQ) test (Gabrielsson, Schenkman, & Hagerman, 1988), where the endpoints 0 and 10 corresponded to "minimum clarity" and "maximum clarity", respectively. Clarity was chosen as the descriptor for sound quality because it was the primary factor our listeners reported using to judge the sentences they heard in Experiment 1. After each trial, listeners were given visual feedback indicating the correct VC.

RESULTS

Experiment 1

Results from the listeners' performance on the sentence quality preference task are reported as the percent of times a listener preferred a specific SPC strength (Fig. 4). Selection rate has been shown to be a valid manner of analysis in a paired-comparison task (Eisenberg, Dirks, & Gornbein, 1997). As SPC strength increased normal-hearing listeners' preference scores decreased, showing a preference for the unprocessed sentences over the SPC-processed sentences. This same pattern was observed in only one of the nine listeners with hearing loss. Six listeners with hearing loss showed little difference between their preference for unprocessed and minimally processed stimuli. The two listeners whose PTAs were 41 and 75 dBHL preferred 1.1 and 1.3 SPC processed sentences, respectively. These results suggest that listeners with more hearing loss prefer stronger SPC strengths. It should be noted that PTA was calculated based on the average of a listener's hearing thresholds at .5, 1, 2, and 4 kHz. There was a significant positive correlation between listeners' PTAs and preferred correction strength ($r = 0.894$, $p = 0.0164$). However, the correlation between PTA and correction strength was not significant when the listener with severe hearing loss (PTA = 75 dB HL) was removed from the analysis. Given this limited set of listeners it is difficult to make any strong conclusion about the relationship between degree of hearing loss and preferred SPC strength, but the results are suggestive.

INSERT FIGURE 4 HERE.

Listeners were asked to describe the basis for their judgments. All listeners reported that the clarity of the stimuli determined their preferences. Clarity has been reported previously as the most significant factor in determining overall sound quality and hearing aid satisfaction (e.g., Gabrielsson et al., 1988; Preminger & Van Tasell, 1995; Eisenberg et al., 1997; Keidser, Dillon, Silberstein, & O'Brien, 2003). Some listeners also reported that their preference for certain stimuli was related to the "fullness" and/or "loudness" of the sound.

Experiment 2

Listeners' clarity ratings of the VC stimuli on a ten-point scale are shown in Fig. 5. Clarity ratings for two normal-hearing listeners decreased monotonically as SPC strength increased, which is similar to how the normal-hearing listeners judged the quality of the sentences in Experiment 1. The other two normal-hearing listeners judged the clarity of the VCs to be the same across all five SPC strengths. No difference in clarity ratings across SPC strengths was observed by four of the five listeners with hearing loss. However, normal-hearing listeners' overall clarity ratings of the unprocessed stimuli (SPC = 1.0) were higher than for listeners with hearing loss. VC clarity ratings for the youngest listener (24 years old) in this study had clarity ratings that decreased as SPC strength was increased. Interestingly, this listener's overall percent correct VC recognition score was more similar to the normal-hearing listeners' scores than to the listeners with hearing loss.

INSERT FIGURE 5 HERE.

The individual phoneme scores for Listener NH-2 and HI-4, shown in Fig. 6, are typical of those obtained by the normal-hearing listeners and listeners with hearing loss, respectively. The asterisks indicate phonemes that were correctly identified more often with SPC processing than without. Normal-hearing listeners obtained high recognition scores for all 16 phonemes in the uncorrected condition. This ceiling effect might be why there were little to no improvements in scores for the SPC conditions. However, the SPC processing did not decrease normal-hearing listeners' overall recognition scores. For HI-4, the listener with hearing loss, SPC improved the scores for phonemes /p/, /t/,/θ/, /z/, and /n/) by more than 10 - 30%. Other phonemes scores (e.g., /s/ and /ð/) were barely above the level of chance (i.e., 6.25%). No single correction strength improved the recognition of all phonemes.

INSERT FIGURE 6 HERE.

Overall percent correct recognition scores were transformed to rationalized arcsine units (RAU) to stabilize variance (Studebaker, 1985). Phoneme recognition scores in RAU, collapsed across all phonemes, are shown in Fig. 7. Normal-hearing listeners scored over 90% regardless of SPC strength, whereas only one listener with hearing loss performed above 70% for any SPC strength. This listener was the youngest listener (24 years old) who has worn binaural hearing aids since pre-school. Although the differences in percent correct scores across different SPC strengths are small, several listeners with hearing loss obtained their highest recognition score with SPC strengths of 1.15 or 1.225.

There was no significant correlation between PTA of 500, 1000 and 2000 Hz for listeners with hearing loss and the SPC strength that yielded their highest overall recognition score in RAU ($r = 0.560$, $p = 0.326$). Again, the range of PTAs for this group of listeners with hearing loss was limited (i.e., 36.7 – 53.8 dB HL).

INSERT FIGURE 7 HERE.

Confusion matrices of listeners' errors on the VC intelligibility test were subjected to Sequential Information Analysis (SINFA) (Wang & Bilger, 1973). The proportion of information transmitted for the acoustic features, including voicing, place and manner, are reported in Table 2. For most subjects the percent of information transmitted remained unchanged or was slightly higher with some level of SPC correction. Two exceptions included HI-9, who showed a large increase in voicing information transmitted at the 1.25 SPC strength, and HI-6, who showed a large decrease in manner information transmitted at the 1.3 SPC strength. These findings suggest that SPC processing does not have any one systematic effect on the main features of speech, but could have a more global effect on phoneme perception.

INSERT TABLE 2 HERE.

Given the large variability in SPC performance observed across listeners with hearing loss, test-retest reliability was examined for one listener with hearing loss. This listener was randomly selected and retested on the same protocol four months after the listener's original test. A simple correlation test indicated good repeatability across sessions in both quality rating ($r = 0.903$, $p < 0.001$) and phoneme recognition ($r = 0.907$, $p < 0.001$).

DISCUSSION

A physiologically-based signal-processing strategy, SPC, was described in this study as a potential new approach to enhance recognition and perceived quality of speech in listeners with hearing loss. SPC introduces different delays across frequency channels of a signal in an attempt to "correct" the abnormal spatiotemporal response pattern of the impaired ear without changing the magnitude spectrum of the sound. Results from the current study showed that SPC improved the sound quality of sentences for most listeners with moderate hearing loss while retaining and in some cases improving the intelligibility of phonemes. Normal-hearing listeners and listeners with mild hearing loss tended to prefer the unprocessed sentences.

Normal-hearing listeners' performance on the preference task in Experiment 1 differed from the normal-hearing listeners' clarity ratings in Experiment 1. These differences could be attributed to the test paradigm and stimuli that were used. For example, in Experiment 1 listener's judgments of sentence quality were obtained using a 2-AFC task, while in Experiment 2 a categorical rating scale was used to judge the clarity of nonsense syllables. A categorical scale might not have been sensitive enough to measure small changes in phoneme clarity, especially for small differences in SPC strengths. Eisenberg et al. (1997) demonstrated that clarity judgments based on a categorical rating system are less sensitive than a paired-comparison scheme, at least for listeners with hearing loss. In addition, neither sentences nor NST are the ideal stimuli. Continuous discourse has been reported to be the most appropriate stimulus in a quality-rating task for speech (e.g., Stelmachowicz, Lewis, & Carney, 1994, Preminger, Neumann, Bake, Walters, & Leavitt, 2000), but cannot be used in an SPC experiment until the speech signal can be SPC processed in real time. However, one advantage of using NST stimuli is that it allowed us to analyze the specific types of improvements and errors related to the SPC processing.

A ceiling effect was observed for the normal- hearing listeners' performance on the VC recognition task. Although this precluded the observation of any considerable improvements in phoneme recognition scores, it cannot explain the lack of any decline in performance as SPC strength increased. It was somewhat surprising that adding the temporal distortions to a normal ear did not have a more negative impact on the normal hearing listeners' recognition scores. Most listeners with hearing loss showed some improvement in their processed recognition scores compared to their unprocessed scores. The degree of this improvement was small. However, the SPC strategy was only applied to frequencies below 2000 Hz and many of the listeners who participated in this study had more hearing loss in the higher than lower frequencies.

Although listeners who benefited the most from SPC had a relatively flat hearing loss, listeners with high-frequency hearing loss also received some benefit from the SPC. There is evidence that a high-frequency hearing loss does influence low-frequency perception of speech (Horwitz, Dubno, & Ahlstrom, 2002). In fact, Doherty & Lutfi (1997) reported that listeners with high-frequency sloping sensorineural loss had difficulty weighting low-frequency components of a complex signal in a selective listening task. Thus, signal-processing schemes targeted at low frequencies may still bring benefit to listeners with hearing loss, regardless of the configuration of their loss.

Interestingly, based on SINFA analysis, SPC did not consistently improve any single acoustic feature of speech. We predicted that the improvement in phoneme recognition would have been associated with an enhancement in some speech cues that would result in a consistent improvement in specific phonemes. However, the improvements and decline in phoneme recognition varied across listeners. Because SPC was not applied to frequencies above 2000 Hz, its effect on speech cues such as noise bursts for plosive identification and frication noise for fricative identification is limited. SPC might have a greater effect on other speech cues such as formant transitions, which are more predominant in low to mid frequencies. Formant transitions are essential for correct identification of plosives (e.g., Kewley-Port, 1983; Dorman, Marton, Hannley, & Lindholm, 1985), fricatives (e.g., LaRiviere, Winitz, & Herriman, 1975; Gelfand, Piper, & Silman, 1986; Pittman & Stelmachowicz, 2000; Nittrouer, 2002) and nasals (e.g., Kurowski & Blumstein, 1984; Qi & Fox, 1992). Future experiments should include a larger set of speech stimuli to help identify which acoustic cues that are most affected by SPC.

One of the challenges in the practical application of SPC is to estimate the loss of nonlinear properties in the impaired ear in an effort to identify the specific SPC strength that would maximally compensate for a given loss. Keep in mind that this loss is not equivalent to audiometric hearing loss. The loss in group delay in an impaired ear could signify other pathologies related to the loss of nonlinearity. In this study, albeit a small group of listeners, severity of hearing loss only served as a modest indicator of preferred correction strength. A larger study with groups of subjects having a range of PTAs from mild to severe is needed to assess the relationship between PTA and SPC strength. To avoid SPC strengths being arbitrarily selected, as was done in the current study, a real-time adjustable SPC "tuner" would be the method of choice to determine a listener's most appropriate correction strength. Speech recognition scores and quality ratings would likely improve with better control over the SPC strength selected for individual listeners. Because group delay is closely associated with cochlear nonlinearity (e.g., Carney, 1994; Cheathum & Dallos, 1998), another way to reach the optimal SPC strength for a specific hearing loss is to explore the relationship between group delay and cochlear biomechanics. For example, otoacoustic emissions (OAEs) are an indirect measure of cochlear nonlinearity (e.g., Brownell, 1990; Neely, Gorga, & Dorn, 2003). Deeper insight might be gained by investigating the connection between OAEs and listeners' preferred and most beneficial SPC strengths. However, a change in group delay is only one aspect of the healthy nonlinear cochlear. Future studies will explore this aspect in more detail as well as other aspects of the cochlear response.

ACKNOWLEDGMENTS

This work was supported by Grant R21 DC006057 from the National Institute on Deafness and Other Communication Disorders. The authors would like to thank Michael Anzalone for his great help in programming and stimulus preparation. The authors would also like to thank Lauren Calandruccio for her effort in collecting part of the data in Experiment 1. Portions of the work were presented at the 2004 American Auditory Society meeting in Scottsdale, Arizona and the 147th meeting of the Acoustical Society of America in New York, New York.

REFERENCES

ANSI. (1989). ANSI-S3.6-1989, *American National Standards Specifications for Audiometers*. New York: American National Standards Institute.

Boothroyd, A., Springer, N., Smith, L., & Schulman, J. (1988). Amplitude compression and profound hearing loss. *Journal of Speech and Hearing Research*, *31*, 362-376.

Brownell, W. E. (1990). Outer hair cell electromotility and otoacoustic emissions. *Ear and Hearing*, *11*, 82-92.

Carney, L. H. (1994). Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness. *Hearing Research*, *76*, 31-44.

Carney, L. H., & Yin, T. C. T. (1988). Temporal coding of resonances by low frequency auditory nerve fibers: Single fiber responses and a population model. *Journal of Neurophysiology*, *60*, 1653-1677.

Cheathum, M. A., & Dallos, P. (1998). The level dependence of response phase: Observations from cochlear hair cells. *Journal of the Acoustic Society of America*, *104*, 356-369.

Dallos, P., & Harris, D. (1978). Properties of auditory nerve responses in absence of outer hair cells. *Journal of Neurophysiology*, *41*, 365-383.

Doherty, K. A., & Lutfi, R. A. (1997). Level discrimination of single tones in a multitone complex by normal-hearing and hearing-impaired listeners. *Journal of the Acoustic Society of America*, *105*, 1831-1840.

Dorman, M. F., Marton, K., Hannley, M. I., & Lindholm, J. M. (1985). Phonetic identification by elderly normal and hearing-impaired listeners. *Journal of the Acoustic Society of America*, *77*, 664-670.

Dubno, J. R., & Dirks, D. D. (1989). Auditory filter characteristics and consonant recognition for hearing-impaired listeners. *Journal of the Acoustic Society of America*, *85*, 1666-1675.

Dubno, J. R., & Shaefer, A. B. (1995). Frequency selectivity and consonant recognition for hearing-impaired listeners with equivalent masked thresholds. *Journal of the Acoustic Society of America*, *97*, 1165-1174.

Eisenberg, L. S., Dirks, D. D., & Gornbein, J. A. (1997). Subjective judgments of speech clarity measured by paired comparisons and category rating. *Ear and Hearing*, *18*, 294-306.

Florentine, M., Buus, S., Scharf, B., & Zwicker, E. (1980). Frequency selectivity in normally-hearing and hearing-impaired observers. *Journal of Speech and Hearing Research*, *23*, 646-649.

Gabrielsson, A., Schenkman, B. N., & Hagerman, B. (1988). The effects of different frequency responses on sound quality judgments and speech intelligibility. *Journal of Speech and Hearing Research*, *31*, 166-177.

Gelfand, S. A., Piper, N., & Silman, S. (1986). Consonant recognition in quiet and in noise with aging among normal hearing listeners. *Journal of the Acoustic Society of America*, *80*, 1589-1598.

Hedrick, M. S., & Rice, T. (2000). Effect of a single-channel wide dynamic range compression circuit on perception of stop consonant place of articulation. *Journal of Speech, Language, and Hearing Research*, *43*, 1174-1184.

Heinz, M. G., Zhang, X., Bruce, I. C., & Carney, L. H. (2001). Auditory-nerve model for predicting performance limits of normal and impaired listeners. *Acoustics Research Letters Online*, *2*, 91-96.

Hickson, L., Thyer, N., & Bates, D. (1999). Acoustic analysis of speech through a hearing aid: Consonant-vowel ratio effects with two-channel compression amplification. *Journal of the American Academy of Audiology*, *10*, 549-556.

Hohmann, V. (2002). Frequency analysis and synthesis using a Gammatone filterbank. *Acta Acustica United with Acustica*, *88*, 433–442.

Horwitz, A. R., Dubno, J. R., & Ahlstrom, J. B. (2002). Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing. *Journal of the Acoustic Society of America*, *111*, 409-416.

Keidser, G., Dillon, H., Silberstein, H., & O'Brien, A. (2003). Sound quality in hearing aids. In *National Acoustic Laboratories Research & Development Annual Report 2000/2003*, Sydney, 40-42.

Kewley-Port, D. (1983). Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustic Society of America*, *73*, 322-335.

Kurowski, K., & Blumstein, S. E. (1984). Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *Journal of the Acoustic Society of America*, *76*, 383-390.

LaRiviere, C., Winitz H., & Herriman, E. (1975). The distribution of perceptual cues in English prevocalic fricatives. *Journal of Speech and Hearing Research*, *18*, 613-622.

Leek, M. R., & Summers, V. (1993). Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise. *Journal of the Acoustic Society of America*, *94*, 3127-3137.

Levitt, H., & Resnick, S. B. (1978). Speech reception by the hearing-impaired: Methods of testing and the development of new tests. In C. Ludvigsen and J. Barfod (Eds.), *Scandinavian Audiology Suppliment*, *6*, 107-129.

Lippman, R., Braida, L., & Durlach, N. (1981). Study of multichannel amplitude compression and linear amplification for persons with sensorineural hearing loss. *Journal of the Acoustic Society of America*, *69*, 524-534.

Moore, B. C. J. (1985). Frequency selectivity and temporal resolution in normal and hearing-impaired listeners. *British Journal Audiology*, *19*, 189-201.

Moore, B. C. J., Vickers, D. A., Plack, C. J., & Oxenham, A. J. (1999). Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *Journal of the Acoustic Society of America*, *106*, 2261-2278.

Neely, S. T., Gorga, M. P., & Dorn, P. A. (2003). Cochlear compression estimates from measurements of distortion-product otoacoustic emissions. *Journal of the Acoustic Society of America, 114*, 1499-1507.

Nelson, D. A. (1991). High-level psychophysical tuning curves: Forward masking in normal-hearing and hearing-impaired listeners. *Journal of Speech and Hearing Research*, *34*, 1233-1249.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and noise. *Journal of the Acoustic Society of America*, *95*, 1085-1099.

Nittrouer, S. (2002). Learning to perceive speech: How fricative perception changes, and how it stays the same. *Journal of the Acoustic Society of America*, *112*, 711-719.

Oxenham, A. J., & Bacon, S. P. (2003). Cochlear compression: Perceptual measures and implications for normal and impaired hearing. *Ear and Hearing*, *24*, 352-366.

Pittman, A. L., & Stelmachowicz, P. G. (2000). Perception of voiceless fricatives by normal-hearing and hearing-impaired children and adults. *Journal of Speech and Hearing Research*, *43*, 1389-1401.

Preminger, J. E., Neuman, A. C., Bakke, M. H., Walters, D. & Levitt, H. (2000). An examination of the practicality of the simplex procedure. *Ear and Hearing*, *21*, 177-193.

Preminger, J. E., & Van Tasell, D. J. (1995). Measurement of speech quality as a tool to optimize the fitting of a hearing aid. *Journal of Speech and Hearing Research*, *38*, 726-736.

Preminger, J. E., & Wiley, T. L. (1985). Frequency selectivity and consonant intelligibility in sensorineural hearing loss. *Journal of Speech and Hearing Research*, *28*, 197-206.

Qi, Y., & Fox, R. A. (1992). Analysis of nasal consonants using perceptual linear prediction. *Journal of the Acoustic Society of America*, *91*, 1718-1726.

Richie, C., Kewley-Port, D., & Coughlin, M. (2003). Discrimination and identification of vowels by young, hearing-impaired adults. *Journal of the Acoustic Society of America*, *114*, 2923-2933.

Souza, P. E., & Bishop, R. D. (1999). Improving speech audibility with wide dynamic range compression in listeners with severe sensorineural loss. *Ear and Hearing*, *20*, 461-470.

Souza, P. E., & Kitch, V. (2001). The contribution of amplitude envelope cues to sentence identification in young and aged listeners. *Ear and Hearing*, *22*, 112-119.

Souza, P. E., & Turner, C. W. (1999). Quantifying the contribution of audibility to recognition of compression-amplified speech. *Ear and Hearing*, *20*, 12-20.

Stelmachowicz, P. G., Kopun, J., Mace, A., Lewis, D. E., & Nittrouer, S. (1995). The perception of amplified speech by listeners with hearing loss: Acoustic correlates. *Journal of the Acoustic Society of America*, *98*, 1388-1399.

Stelmachowicz, P. G., Lewis, D. E., & Carney, E. (1994). Preferred hearing aid-frequency responses in simulated listening environments. *Journal of Speech and Hearing Research*, *37*, 712-719.

Stone, M. A., & Moore, B. C. J. (1992). Spectral feature enhancement for people with sensorineural hearing impairment: Effects on speech intelligibility and quality. *Journal of Rehabilitation Research and Development*, *29*, 39-56.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research*, *28*, 455-462.

Studebaker, G. A., & Sherbecoe, R. L. (1995). Speech recognition at higher than normal speech and noise levels. *Journal of the Acoustic Society of America*, *97*, 3358.

Summerfield, Q. (1987). Speech perception in normal and impaired hearing. *British Medicine Bulletin*, *43*, 909-925.

Turner, C. W., Chi, S., & Flock, S. (1999). Limiting spectral resolution in speech for listeners with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *42*, 773-784.

Turner, C. W., & Henn, C. C. (1989). The relation between vowel recognition and measures of frequency resolution. *Journal of Speech and Hearing Research*, *32*, 49-58.

Van Tasell, D. J. (1993). Hearing loss, speech, and hearing aids. *Journal of Speech and Hearing Research*, *36*, 228-244.

Van Tasell, D. J., & Trine, T. D. (1996). Effects of single-band syllabic amplitude compression on temporal speech information in nonsense syllables and in sentences. *Journal of Speech and Hearing Research*, *39*, 912-922.

Wang, M., & Bilger, R. (1973). Consonant confusions in noise: A study of perceptual features. *Journal of the Acoustic Society of America*, *45*, 1248-1266.

Table 1. Pure-tone air conduction thresholds in dB HL for 6 normal-hearing listeners (NH) and 12 listeners with hearing loss (HI).

Listener		**Frequency (Hz)**								
		250	**500**	**1000**	**1500**	**2000**	**3000**	**4000**	**6000**	**8000**
NH-1	R	5	-5	10		15	5	15	35	30
	L	5	0	10		0	5	15	25	25
NH-2	R	0	0	10		5	0	0	5	10
	L	0	0	0		0	5	5	10	5
NH-3	R	5	5	5		0	15	15	25	15
	L	5	5	5		5	5	10	25	15
NH-4	R	15	5	15		-5	-5	5	5	5
	L	15	5	5		0	-5	-5	-5	10
NH-5	R	20	15	15		0	5	10	10	10
	L	10	10	15		0	5	15	5	10
NH-6	R	5	5	5		10		5	0	5
	L	10	10	10		10		5	5	5
*HI-1	R	20	20	45		55/30	55/30	70/35	75	75
	L	25/0	15	35		25	35	35	55	65
*HI-2	R	40/5	60/15	70/25		80/45		90/75		NR
	L	95	NR	NR		NR		NR		NR
HI-3	R	75	70	60		50	45	55	65	70
	L	20	20	35		40	45	55	70	85
HI-4	R	45	50	55		70	65	75	70	80
	L	45	50	65		80	65	65	70	75
HI-5	R	30	25	25		55	60	65	100	90
	L	20	25	30		55	60	70	90	85
HI-6	R	30	25	30		55	50	55	65	70
	L	35	35	40		55	60	60	80	75
HI-7	R	30	30	50		45	35	50	45	55
	L	20	30	45		40	40	40	50	75
HI-8	R	55	45	50		45	40	50	75	80
	L	45	30	45		50	60	70	75	70
HI-9	R	50	45	50		45	40	50	50	80
	L	50	45	55		55	50	50	65	70
HI-10	R	25	15	15		25		35	55	65
	L	15	20	15		30		40	55	60
HI-11	R	20	15	20	20	45		50	40	50
	L	15	10	15	30	55		60	55	60
HI-12	R	10	10	15	15	40		45	45	50
	L	10	10	20	25	50		45	60	60

* Listeners HI-1 and HI-2 have a mixed hearing loss. Air conduction (AC) and bone conduction (BC) thresholds are displayed as AC/BC. NR refers to "no response" at the limits of the GSI-16 audiometer (105 dB HL).

Table 2. Results from SINFA analysis for listeners with normal hearing (NH) and listeners with hearing loss (HI) on a VC recognition task performed at five different SPC strengths.

Voicing

	SPC	NH-2	NH-3	NH-4	NH-5	HI-4	HI-6	HI-7	HI-8	HI-9
Information Transmitted	1.000	0.884	0.797	0.759	0.838	0.838	0.861	0.967	0.863	0.554
	1.075	0.887	0.762	0.783	0.933	0.741	0.797	0.940	0.839	0.598
	1.150	0.966	0.823	0.789	0.917	0.901	0.860	0.967	0.805	0.575
	1.225	0.967	0.797	0.751	0.907	0.818	0.863	0.943	0.782	0.650
	1.300	0.823	0.800	0.751	0.860	0.966	0.800	1.000	0.875	0.618

Place

	SPC	NH-2	NH-3	NH-4	NH-5	HI-4	HI-6	HI-7	HI-8	HI-9
Information Transmitted	1.000	0.918	0.971	0.939	0.917	0.376	0.550	0.766	0.511	0.450
	1.075	0.921	0.918	0.885	0.962	0.401	0.517	0.745	0.554	0.460
	1.150	0.954	0.950	0.918	0.966	0.353	0.555	0.690	0.548	0.505
	1.225	0.965	0.965	0.918	0.935	0.446	0.517	0.775	0.532	0.747
	1.300	0.945	0.886	0.921	0.933	0.393	0.487	0.755	0.527	0.453

Manner

	SPC	NH-2	NH-3	NH-4	NH-5	HI-4	HI-6	HI-7	HI-8	HI-9
Information Transmitted	1.000	0.903	0.987	0.948	0.921	0.618	0.796	0.981	0.725	0.742
	1.075	0.913	0.962	0.923	0.979	0.607	0.780	0.967	0.782	0.703
	1.150	0.916	0.981	0.913	0.985	0.603	0.825	0.985	0.775	0.709
	1.225	0.981	1.000	0.943	0.919	0.658	0.713	1.000	0.754	0.714
	1.300	0.879	0.928	0.935	0.952	0.707	0.160	1.000	0.823	0.661

FIGURE LEGEND

Figure 1. Schematic illustration of level-dependent changes in both magnitude and phase properties of peripheral filters. Solid lines represent filter properties at high SPLs, and dashed lines represent low SPLs. The gain and bandwidth vary more with level in the normal ear than in the impaired ear. Similarly, changes in the phase properties of the filter vary more as a function of sound level in the normal ear than in the impaired.

Figure 2. Illustration of the relationship between group delay and phase properties of the cochlear filter. Left: Impulse responses of filters in the normal (top panel) and impaired (bottom panel) periphery. The duration of the build-up of the filter's response depends upon how sharply tuned the filter is (filter functions shown at the right). Broad filters have short build-up times, whereas sharp filters have a long build-up time. The build-up time is proportional to the group delay; the vertical lines show the group delay approximation for gammatone filters used in the SPC system. In the normal ear, the actual group delay constantly fluctuates between the low- and high-SPL group-delay values (see arrow labeled dynamic group delay). In the impaired ear, the group delay varies much less across SPLs (vertical lines are closer to each other). However, by adding a dynamic delay (the correction), the normal dynamic group delay can be approximated on the output of the impaired filter.

Figure 3. Schematic diagram of low-frequency SPC system. The control pathways (left) computed the amount of correction in phase delay and then submitted it to the analysis-synthesis filterbank (right).

Figure 4. Preference for SPC strength for 9 listeners with hearing loss. The percentage of times that sentences with each SPC strength were preferred in pair-wise tests is plotted as a function of SPC strength. The bold solid lines (repeated in all three panels) are average preferences for three normal-hearing listeners. The three panels show results for three groups of listeners with hearing loss. Top: Four listeners with hearing loss preferred uncorrected stimuli (SPC strength = 1.0). Middle: Four listeners with hearing loss preferred corrected stimuli with low SPC strengths (1.05 -1.1). Bottom: One listener with severe hearing loss preferred a high SPC strength (1.25). PTAs (500, 1000, 2000, and 4000 Hz) are shown for each listener in the legends.

Figure 5. Clarity rating as a function of correction for 16 NST VCs in 4 normal-hearing listeners (NH, upper panel) and 5 listeners with hearing loss (HI, lower panel). VCs differed in the ending-consonant phonemes. Presentation level was fixed at each listener's MCL. Each line with a different symbol represents the data from one listener. Data were averaged across 16 VCs.

Figure 6. Phoneme-recognition scores in one normal-hearing listener (NH-2, left panel) and one listener with hearing loss (HI-4, right panel). Each vertical bar within a cluster of 5 bars represents one recognition score for a specific phoneme. Each set of bars shows scores for SPC strengths varying from 1.0 (uncorrected) to 1.3, from left to right. Each bar represents the results for 16 trials at a given stimulus condition. The legend shows the correction strengths corresponding to the bars of different shades.

Figure 7. Phoneme recognition in RAU as a function of correction strength in 4 normal-hearing listeners (NH) and 5 listeners with hearing loss (HI). Each line with a different symbol represents the data from one listener. Arrows bracket the results for each group of listeners. Data were averaged across 16 phonemes.

FIGURE 1

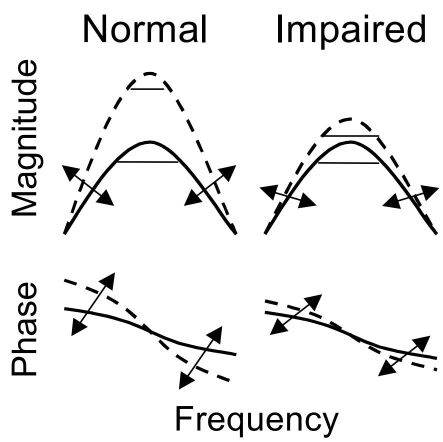
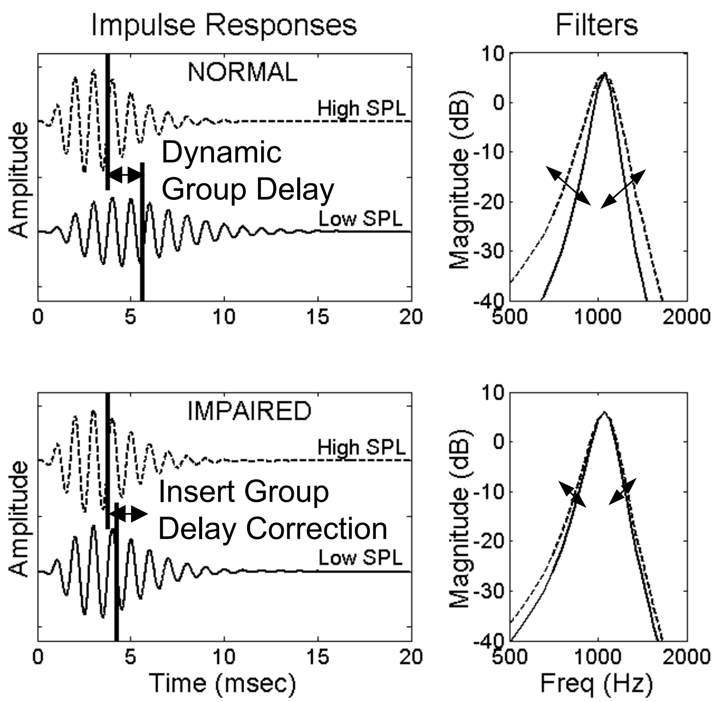
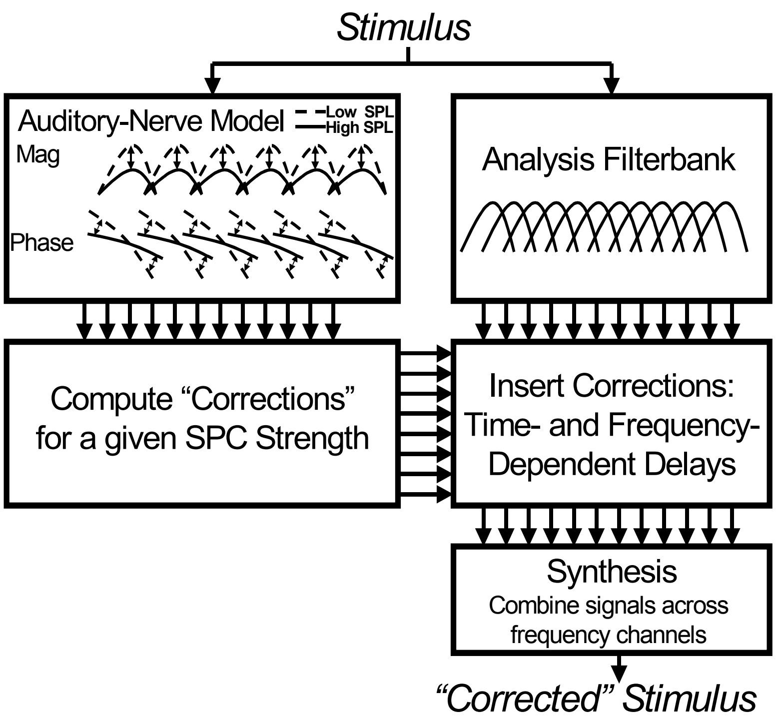


FIGURE 2



Impulse Responses
Filters
NORMAL
High SPL
Dynamic
Group Delay
Low SPL
Amplitude
Magnitude (dB)
IMPAIRED
High SPL
Insert Group
Delay Correction
Low SPL
Amplitude
Magnitude (dB)
Time (msec)
Freq (Hz)

FIGURE 3



Stimulus
Auditory-Nerve Model
Low SPL
High SPL
Mag
Phase
Analysis Filterbank
Compute “Corrections” for a given SPC Strength
Insert Corrections: Time- and Frequency-Dependent Delays
Synthesis
Combine signals across frequency channels
“Corrected” Stimulus

FIGURE 4

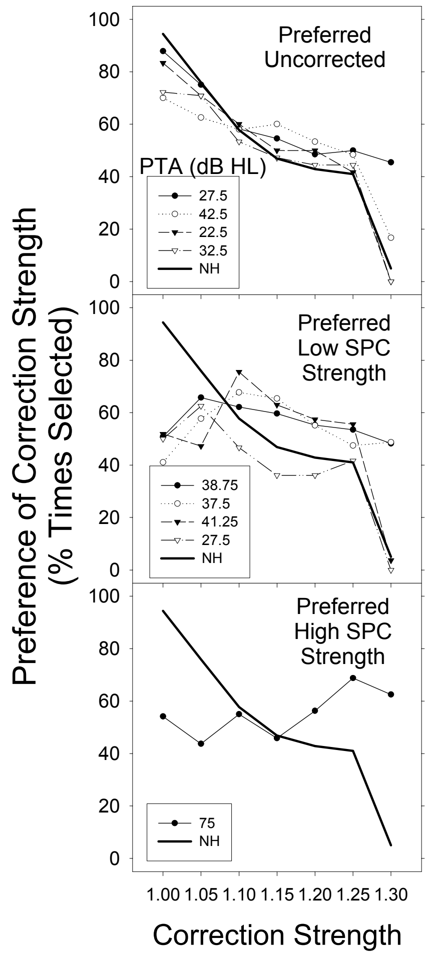

Preferred
Uncorrected
PTA (dB HL)
27.5
42.5
22.5
32.5
NH
Preferred
Low SPC
Strength
38.75
37.5
41.25
27.5
NH
Preferred
High SPC
Strength
75
NH
Preference of Correction Strength
(% Times Selected)
0
20
40
60
80
100
1.00 1.05 1.10 1.15 1.20 1.25 1.30
Correction Strength

FIGURE 5

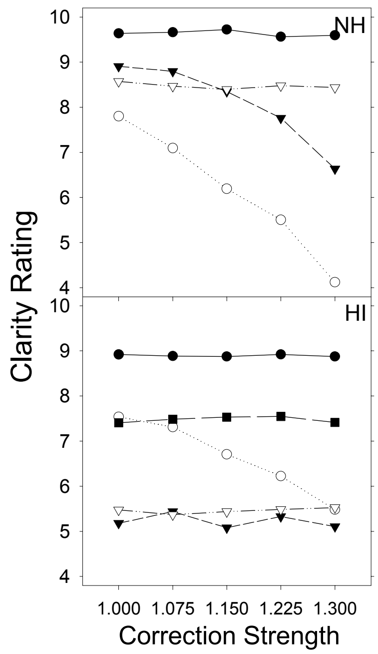

NH
HI
Clarity Rating
10
9
8
7
6
5
4
1.000
1.075
1.150
1.225
1.300
Correction Strength

FIGURE 6

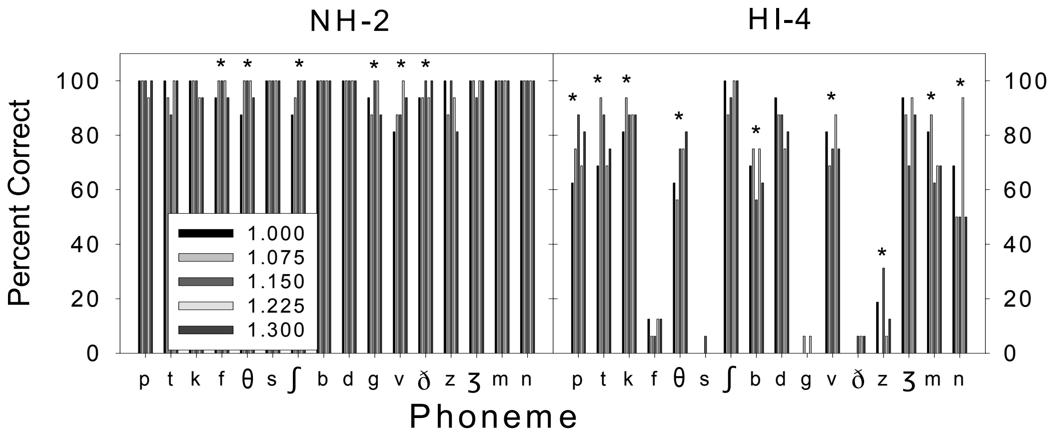
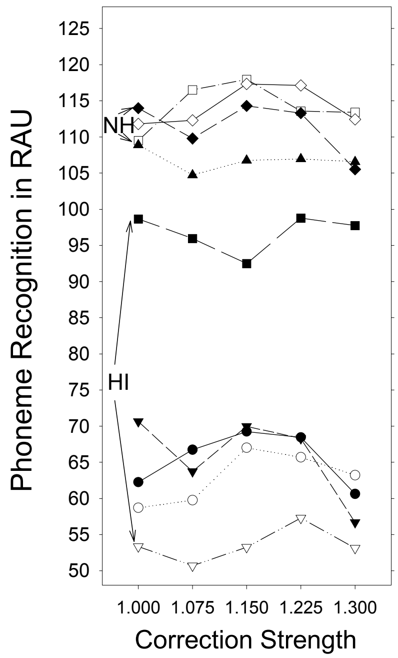

NH-2
HI-4
Percent Correct
0
20
40
60
80
100
1.000
1.075
1.150
1.225
1.300
p t k f θ s ʃ b d g v ð z ʒ m n
Phoneme

FIGURE 7



Phoneme Recognition in RAU
125
120
115
110
105
100
95
90
85
80
75
70
65
60
55
50
NH
HI
1.000
1.075
1.150
1.225
1.300
Correction Strength