Published in final edited form as:

J Speech Lang Hear Res. 2007 June ; 50(3): 554–563. doi:10.1044/1092-4388(2007/038).

Temporal intra-speech masking of plosive bursts: Effects of hearing loss and frequency shaping

Carol L. Mackersie

School of Speech, Language and Hearing Sciences San Diego State University

Abstract

PURPOSE—The purposes were 1) to compare masking of consonant bursts by adjacent vowels for listeners with and without hearing loss and 2) to determine the extent to which the “temporal intra-speech masking” can be reduced by a simulated hearing aid frequency-response shaping.

METHOD—Fourteen adults with sensorineural hearing loss and six with normal hearing served as participants. Seven of the participants with hearing loss had flat/gradually sloping audiograms and seven had steeply sloping losses. Stimuli consisted of three consonant bursts (/t/, /p/, /k/) presented in isolation and in vowel-consonant-vowel combinations using the vowel /a/ with formant transition information removed. Normal-hearing listeners were tested using unfiltered stimuli. Listeners with hearing loss were tested using unfiltered stimuli and stimuli filtered to approximate a hearing aid frequency response prescribed by NAL-R. All listeners were tested under earphones at the most comfortable level for the vowel stimulus. Temporal intra-speech masking was quantified as the threshold shift produced by the adjacent vowels.

RESULTS—Average intra-speech masking for listeners with steeply sloping hearing loss was significantly higher than that of normal-hearing listeners and those with flat/gradually sloping losses. Greater intra-speech masking was observed for /t/ and /p/ than for /k/. On average, frequency shaping significantly reduced the amount of intra-speech masking for listeners with steeply sloping hearing losses. Even with appropriate amplification/spectral shaping, however, temporal intra-speech masking remained greater than normal for several individuals.

CONCLUSIONS—Finding suggest that some individuals with steeply sloping losses may need additional signal processing to supplement frequency-shaping in order to overcome the effect of temporal intra-speech masking.

Intra-speech masking refers to the interference in the perception of speech cues by components of the speech signal itself. Intra-speech masking has typically been examined using synthesized vowels by measuring the perception of the second formant (F2) with reductions in the level of the first formant (F1) (Danaher et al., 1973; Danaher & Pickett, 1975; Pickett & Danaher, 1975; Van Tasell, 1980; Dorman et al., 1985a; Summers & Leek, 1997). Several of these studies demonstrated that attenuation of F1 can improve the audibility and discriminability of F2 transitions (Hannley & Dorman, 1983; Pickett & Danaher, 1975; Summers & Leek, 1997; Van Tasell, 1980).

The studies noted above examined simultaneous intra-speech masking, that is, masking that occurs within simultaneous components of a speech signal. There is some evidence to suggest, however, that temporal masking between sequential speech cues can also influence perception. A number of studies have examined the influence of silence duration on the perception of intervocalic stop consonants in listeners with and without sensorineural hearing loss. Generally, as the silence duration is decreased, listeners become less accurate at labeling plosives on a voicing continuum (Lisker, 1957). Some listeners with hearing loss require abnormally long silence duration to accurately perceive voicing of plosives when consonants occur in final or

medial positions compared to listeners with normal hearing (Cazals & Palis, 1991; Cazals, 1994; Dorman et al., 1985b). Longer recovery time from forward masking for listeners with hearing loss has been proposed as an explanation for these differences between listeners with and without hearing loss (Cazals & Palis, 1991).

Revoile et al. (1981) examined temporal intra-speech masking in listeners with sensorineural hearing loss by measuring the masking of simulated consonant bursts by adjacent synthetic vowels. Consonant bursts were simulated using 50 ms band-limited noises of 500-1500 Hz, 1500-4000 Hz, and 4000-6000 Hz. The maskers were two-formant synthetic vowels /a/ and /i/. Detection thresholds for the simulated consonant bursts were measured in isolation and in the presence of the vowels using separations of 10 and 100 ms between the vowels and bursts. Separate estimates were obtained for vowels preceding and succeeding the bursts to estimate the magnitude of both forward and backward masking. Masking was greatest in the vowel +burst condition (forward masking) for the low-frequency noise burst masked by /a/, reflecting the condition with the greatest spectral overlap. At the 100 ms temporal separation, most closely resembling natural speech, the average amount of forward masking was 5 dB or less.

The work of Revoile et al. (1981) suggests that the temporal masking of consonant bursts by adjacent vowels may be negligible for listeners with hearing loss. This finding however, contrasts with the observation described earlier that listeners with hearing loss require greater inter-vocalic silence duration to accurately perceive voicing cues. It is not clear whether the pattern of results obtained by Revoile et al. (1981) would be the same for real speech. Real vowels may have a greater masking effect on consonant bursts because the spectral density of real vowels is considerably greater than the two-formant synthesized vowels used by Revoile and her colleagues. In addition, spectro-temporal characteristics of real consonant bursts are different from the filtered noise bands used in the study by Revoile.

The general goal of the present study was to examine the masking of intervocalic consonant bursts in listeners with normal hearing and sensorineural hearing loss using stimuli derived from real speech. Throughout this paper, the term “temporal intra-speech masking” will be used to refer to the shift in consonant burst detection threshold produced by the adjacent vowels. Consonant bursts were used as stimuli because of their importance in the perception of place cues, particularly for listeners with hearing loss (Cassidy & Harrington, 1995; Raz & Noffsinger, 1985). Raz and Noffsinger (1985), for example, reported that removal of consonant burst cues from labeling continua (/ba/, /da/, /ga/) resulted in a substantial deterioration in performance for participants with hearing loss, but not for listeners with normal hearing. In addition, the levels of consonant bursts have been shown to be better predictors of nonsense syllable recognition than are formant transition cues (Stelmachowicz et al. 1995).

The present study supplements previous work on temporal intra-speech masking by evaluating the effects of hearing loss configuration, and by determining the effects of frequency shaping on the magnitude of masking effects. The study addressed two questions: 1) What are the effects of sensorineural hearing loss and hearing loss configuration on the magnitude of temporal intra-speech masking using stimuli created from natural speech? and 2) To what extent can simulated hearing aid frequency-response shaping reduce the magnitude of temporal intra-speech masking in listeners with hearing loss? It is predicted that listeners with steeply sloping hearing losses will have greater temporal intra-speech masking than listeners with gradually sloping/flat losses. This prediction is based on evidence that listeners with sloping hearing loss are generally more susceptible to the effects of upward spread of masking than listeners with flat losses (Cook et al., 1997; Fabry et al., 1993; Gagné, 1988). Simulated hearing aid frequency response shaping is expected to reduce the amount of temporal intra-speech masking by reducing the lower frequency components of the vowel masker while emphasizing the high-frequency components of the consonant bursts.

METHODS

Participants

Ten adults with normal hearing and 14 adults with sensorineural hearing loss participated in the study. The mean age of listeners with hearing loss was 72 years ranging from 43 to 89 years. The mean age of listeners with normal hearing was 27 years ranging from 22 to 44 years. Participants with normal hearing had octave thresholds of 20 dB HL or less between 250 and 8000 Hz. Seven of the participants with hearing loss had gradually sloping or flat configurations (less than 15 dB/octave between 1 kHz and 4 kHz). The remaining participants with hearing loss had steeply sloping configurations (15 dB / octave or more between 1 and 4 kHz). The mean pure-tone thresholds of the two hearing loss groups are shown in Figure 1.

Stimuli

Three vowel-consonant-vowel combinations (VCV) (/aka/, /ata/, /apa/) were recorded by a female talker. After equalizing the levels of the VCV stimuli, based on the vowel rms levels, the vowels were replaced by steady-state stimuli created by replicating one cycle of the steady-state portion of the vowel. The purpose of this change was to eliminate the formant transition cues to burst presence. The total duration of each VCV stimulus was 610 ms consisting of a 225 ms vowel segment, 100 ms silence, 60 ms consonant burst, and 225 ms vowel segment. The same vowel segment was used for the three consonant bursts. Segments were assembled using SigGen software (Tucker-Davis Technologies, 1997). The peak levels of the vowel and consonant bursts measured through the earphone before further processing were 113 dB peak SPL (/a/), 107 dB peak SPL (/p/), 104 dB peak SPL (/k/), and 91 dB peak SPL (/t/).

The spectra of the vowel and each of the three consonant bursts based on analyses of the digital waveforms are shown in Figure 2. Note that the levels are expressed in dB digital (re: arbitrary zero). The spectra were obtained from 256 point samples using a hanning window. Unfiltered speech spectra are shown in the left panels. Filtered spectra are shown in the right panels for one subject with a gradually sloping hearing loss (6.25 dB / octave). As shown in the figure, filtering produces an increase in the high-frequency portion of the consonant burst spectra relative to the primary peak of the vowel.

Participants with normal hearing heard only the unfiltered stimuli, whereas participants with hearing loss heard both the unfiltered and spectrally shaped stimuli. For the spectrally shaped conditions, stimuli were filtered for each participant to approximate the frequency response prescribed by NAL-R (Byrne & Dillon, 1986). Filters were generated using Tucker-Davis Technologies (TDT) FIR software and implemented using a TDT programmable filter (PF1). The filtering action was verified using Tucker-Davis XFUNCT software by playing a complex signal through the loaded filter.

Digitized stimuli were played from a Pentium computer using Tucker-Davis System II hardware consisting of a D/A converter (50 kHz sampling rate), a programmable filter (PF1) and a programmable attenuator (PA4). Attenuated signals were routed to a headphone buffer (HB6) and delivered to Sennheiser HD 265 headphones. The presentation level of the signals were individualized to correspond to each participant's most comfortable loudness level as determined by a categorical loudness rating procedure (described below). The choice of most comfortable loudness level was based on a desire to simulate the volume control/gain settings generally preferred by hearing aid users (Byrne & Cotton, 1988).

Procedures

Loudness Measures Categorical loudness ratings were used to estimate the most comfortable loudness levels of the vowels using a procedure described by Hawkins et al. (1987). The

intensity level of the vowel was increased in 5 dB steps starting at or below the subject's pure-tone average. The starting level was varied randomly over a 10 dB range between 0 and 10 dB SL. At each level, subjects were asked to rate the loudness of the sound according to the following categories: "cannot hear", "very soft", "soft", "comfortable, but slightly soft", "comfortable", "comfortable, but slightly loud" and "loud, but ok". This procedure was repeated until participants' performance stabilized. The normal hearing participants rated the unfiltered stimuli only, whereas the participants with hearing loss rated both unfiltered and spectrally-shaped stimuli.

Threshold Procedure The levels of the consonant bursts were manipulated to determine the burst thresholds in isolation and in the VCV context. The threshold of the consonant bursts in isolation provided an estimate of the audibility of the bursts. The threshold elevation produced by adding the adjacent vowels provided an estimate of the temporal intra-speech masking produced by the vowels.

Burst thresholds were estimated using a two-interval forced-choice adaptive procedure incorporating a three-down one-up rule. This procedure estimates the 79.4% correct point on the psychometric function (Levitt, 1971). An initial step size of 5 dB was reduced to the minimum step size of 2 dB following three reversals. Each block of trials was terminated following 12 reversals. Threshold estimates were based on the last eight reversals. A minimum of three threshold estimates were obtained for each condition. Testing continued until 3 consecutive estimates were within 3 dB.

RESULTS

Temporal intra-speech masking without frequency shaping

Figure 3 shows average temporal intra-speech masking for unfiltered stimuli for listeners with steep and flat/gradual hearing loss slopes (bars) and for individual listeners (symbols). The means and upper limits of the 95% confidence intervals for normal hearing listeners are indicated by the solid and broken horizontal lines, respectively. The ordering of data points for individual listeners (left to right) is identical in each panel.

As shown in Figure 3, listeners with hearing loss exhibited more temporal intra-speech masking than did listeners with normal hearing with the greatest amount of masking seen for the listeners with steeply sloping hearing losses. There were substantial intersubject differences, particularly within the group of listeners with steeply sloping hearing losses. A repeated measures analysis of variance (ANOVA) indicated a significant interaction between hearing status and consonant burst reflecting greater differences among listeners groups for /t/ and /p/ than for /k/ [$F(4,36) = 3.09, p < .05$]. Newman-Keuls post-hoc testing indicated significant differences among the consonants for the listeners with steeply sloping losses, but not for the other listeners. Specifically, there was significantly more intra-speech masking of the /t/ and /p/ bursts than for the /k/ burst ($p < .001$), with no significant difference between /t/ and /p/ ($p > .05$). In addition, there were no significant differences between the mean intra-speech masking for listeners with gradually sloping hearing loss and normal hearing, for any of the consonant bursts ($p > .05$). The magnitude of masking for several individuals with flat/gradually sloping losses, however, exceeds the 95% confidence interval of the normal-hearing listeners by a substantial amount.

Individual differences

Substantial individual differences in the amount of intra-speech masking were observed among listeners with hearing loss. To examine the possible factors contributing to individual differences, stepwise forward regression analyses were conducted using listener age, pure-tone

average, and hearing loss slope (dB change between 500 and 4000 Hz) as predictors. Only factors with F -ratios of greater than 1.0 were included in the model. Table 1 shows the results of these analyses completed separately for each consonant burst. In Figure 4, temporal speech masking for individual listeners is plotted as a function of hearing loss slope. Generally, an increase in hearing loss slope is associated with greater intra-speech masking. Results of the regression analyses indicated that hearing loss slope accounted for a significant proportion of the variance for the /t/ burst (30%) with age contributing an additional (non-significant) proportion (7%). For the other two consonant bursts there was weaker evidence that hearing loss slope contributed to the model (/k/ burst: $p < .06$; /p/ burst: $p < .08$). The other factors did not reach the F -ratio criterion to be included in the regression model.

Results are consistent with the idea that listeners with more steeply sloping hearing losses are more susceptible to temporal intra-speech masking. In the next section, the question of whether simple hearing aid frequency response shaping can overcome the effects of intra-speech masking will be considered.

Effects of frequency shaping

The effects of frequency shaping on temporal intra-speech masking is shown in Figure 5. It can be seen that frequency shaping reduced temporal intra-speech masking for listeners with steeply sloping losses, but not for listeners with flat/gradually sloping losses.

A repeated-measures ANOVA was completed using filtering, consonant, and listener group as factors. There was a significant interaction between filter condition and listener group [$F(1,12) = 9.36, p < .01$]. Subsequent Newman-Keuls post-hoc testing revealed a significant reduction ($p < .01$) of speech masking for the steeply sloping group, but not for the gradually sloping group. There was a significant interaction between listener group and consonant [$F(2,24) = 3.53, p < .05$]. Newman-Keuls post-hoc testing revealed significant differences in the amount of masking among the consonants for the steeply sloping group ($p < .05$), whereas there were no significant differences in the amount of masking among consonants for the gradually sloping group ($p > .05$).

The analyses above indicate that frequency shaping reduced the amount of intra-speech masking for listeners with steeply sloping hearing losses, but not for those with gradually sloping losses. Given that the mean intra-speech masking for unfiltered speech was similar for listeners with normal hearing and gradually sloping hearing losses, it is not surprising that the listeners with gradually sloping loss did not show substantial reduction in masking with filtering. Further analyses of data for listeners with steeply sloping hearing losses were completed to determine if the excess temporal intra-speech masking is eliminated by spectral shaping. An additional ANOVA was completed to determine if intra-speech masking after spectral shaping was significantly different from intra-speech masking for normal-hearing listeners. The ANOVA indicated a significant interaction between hearing status and consonant [$F(2,24) = 3.89, p < .01$]. Newman-Keuls post-hoc testing confirmed that for participants with steeply sloping hearing losses, speech masking remained higher than normal only for /p/ ($p < .05$). On average, for listeners with steeply sloping losses, the spectral shaping eliminated the excess intra-speech masking of /t/ and /k/ such that the masking magnitude was no longer significantly greater than normal.

In summary, frequency shaping reduced the amount of intra-speech masking for listeners with steeply sloping hearing losses, but not gradually sloping losses. Average intra-speech masking for /k/ and /t/ was similar to that of normal-hearing listeners after spectral shaping. On average, speech masking for /p/ remained higher than normal after spectral shaping for listeners with steeply sloping losses. It is important to note that even within a listener group, there were substantial intersubject differences in the reduction of temporal intra-speech masking with

spectral-shaping. Even with spectral shaping, half the subjects had temporal intra-speech masking that exceeded the 95% confidence interval of normal-hearing listeners.

Acoustic analyses

The overall presentation levels of the vowels were based on estimates of listeners' MCLs determined separately for the unfiltered and filtered stimuli. This introduces the possibility that the presentation levels were different for the unfiltered and filtered vowels. Differences between overall levels of the unfiltered and spectrally-shaped vowels could potentially affect the magnitude of temporal intra-speech masking to the extent that masking increases with increased level (Gagné, 1988; Summers & Leek, 1997). For this reason, differences between filtered and unfiltered presentation levels were calculated for each listener with hearing loss. In addition, consonant-to-vowel ratios were estimated in the unfiltered and spectrally-shaped conditions.

The top panel of Figure 6 illustrates the average differences between vowel presentation levels for the unfiltered and spectrally-shaped stimuli. The average vowel presentation levels corresponding to MCL were within 2 dB for the unfiltered and filtered stimuli for both the steeply sloping and gradually sloping groups. For most listeners, differences between unfiltered and filtered stimuli were less than 3 dB. However, there were two listeners whose MCLs differed by 9 dB; one whose MCL was higher for the unfiltered stimuli, and one whose MCL was higher for the filtered stimuli. The individual variability in MCL levels was within the normal test-retest variability (8-12 dB) observed for speech MCL measures (e.g. Sammeth et al., 1989). As shown in the figure, consonant burst levels were more affected by spectral shaping than vowel levels. The high frequency emphasis resulted in an increase in the level of the consonant bursts /t/ and /k/ compared to the unfiltered. In contrast, the spectral-shaping reduced the substantial low-frequency energy of the /p/ burst resulting in a decrease in overall level.

Consonant-to-vowel ratios of the unfiltered and spectrally shaped stimuli are shown in the bottom panel of Figure 6. It can be seen that the spectral shaping produced an increase in the consonant-to-vowel ratio for /aka/ and /ata/, attributable to an increase in the consonant levels.

Based on the acoustic analysis the general reduction in temporal intra-speech masking in the spectrally-shaped conditions cannot be attributed to a reduction in the presentation level of the vowel. Rather, the reduction in intra-speech masking for listeners with steeply sloping losses more likely a result of the increased in consonant levels and corresponding consonant to vowel ratios. The increase in average /t/ and /k/ burst levels with spectral shaping corresponds with the reduction in intra-speech masking observed for listeners with steeply sloping hearing loss.

DISCUSSION

Listeners with steeply sloping hearing losses experienced greater temporal intra-speech masking of consonant bursts by adjacent vowels than did listeners with flat-gradually sloping losses and with normal hearing. The greater intra-speech masking observed for listeners with steeply sloping hearing loss is consistent with observations of other investigators showing greater upward spread of masking for individuals with greater hearing loss slope (Gagné, 1988; Summers & Leek, 1997).

The differences between the magnitude of temporal intra-speech masking for the /t/, /p/, and /k/ are compatible with the differences between spectral properties of the bursts and the vowel. As shown in Figure 2 (left panels), there is less spectral overlap between the vowel and the burst for /k/ than for the other two consonants, paralleling the finding that intra-speech masking was lowest for the /k/ burst.

The average temporal intra-speech masking of unfiltered consonant bursts was approximately 8 dB, reflecting considerably more masking than the 2-5 dB of masking reported by Revoile et al. (1981) for silence durations closest to those used in the present study. Differences between stimuli used in the two studies may account for the differences between the two studies. Specifically, the spectral density of the vowel masker used in the present study was greater than the synthesized two-formant vowels used by Revoile and her colleagues. Also, the bandwidth of the consonant bursts was larger than the bandwidth of the narrow-band noises used by Revoile et al. (1981) resulting in greater spectral overlap between the vowel and consonant burst stimuli. In addition, the vowel maskers in the present study consisted of vowel segments placed before and after the consonant bursts (forward + backward masking), whereas Revoile's group measured forward and backward masking separately. Finally, the participants in the present study were considerably older than the college-aged students by Revoile et al. (1981). Based on Gehr and Sommers (1999) report that backward masking is significantly greater in older listeners than younger listeners, listener age may also be a contributing factor in the differences in results for the two studies.

The reduction of temporal intra-speech masking observed for listeners with steeply sloping hearing loss in the spectrally-shaped condition compared to the unfiltered condition is consistent with a reduction in upward spread of masking. The excess masking observed for the /t/ and /k/ burst for participants with steeply sloping losses was, on average, effectively eliminated, but not for all participants. Interestingly, for two participants, thresholds for spectrally-shaped consonant bursts were actually higher than for unfiltered stimuli, suggesting that different frequency regions may have contributed to detection of the bursts in the spectrally-shaped and unfiltered conditions. The least reduction in intra-speech masking was observed to the /p/ burst. This may be explained by the similarities in the spectra for the /p/ and /a/ which may have resulted in a similar amount of on-frequency masking in the 2500-3000 Hz regions in the unfiltered and spectrally shaped conditions.

In the current study, vowel presentation level was based on listeners' MCLs measured independently for the filtered and unfiltered stimuli. The use of MCL rather than a fixed presentation level introduced the possibility that vowel presentation level was different for the filtered and spectrally-shaped stimuli. It is notable that in a study examining the effects of increased consonant-to-vowel-ratio on consonant recognition, Gordon-Salant (1987) reported an influence of presentation level on consonant recognition. Specifically, consonant recognition improved by a small amount (4 percentage points) with a 15 dB increase in level for those with gradually sloping losses, but did not improve for those with steeply sloping hearing losses. This result may be attributed to insufficient audibility of high frequency cues important for consonant recognition by listeners with steeply sloping losses.

It is unlikely, however, that presentation level was an influencing factor in the present study for several reasons. First, the differences between the presentation levels for the unfiltered and spectrally shaped vowels were small - within 3 dB for all but two listeners. Second, recognition and detection are fundamentally different and do not rely on identical cues. Consonant recognition is dependent on access to high frequency cues that should be more influenced by differences in hearing loss configuration than consonant burst detection. Because the consonant bursts are broadband, listeners may rely on energy across a larger range of frequencies for simple detection. Most importantly, the variable of interest in the current study was the threshold shift of the consonant burst produced by the addition of the vowel masker, not the absolute threshold of the burst. Because the thresholds of the unmasked consonant bursts were used as the reference, audibility of the consonant burst was ensured to the extent that the adaptive psychophysical procedure resulted in similar points on the psychometric function in the unfiltered and spectrally shaped conditions.

Any influence of presentation level would most likely appear as a difference in the amount of masking (i.e. threshold shift) resulting from an increase in the level of the vowel masker in either the unfiltered or spectrally-shaped conditions. Greater upward spread of masking for higher masker levels has been demonstrated in a number of studies using pure-tone signals. (e.g. Egan & Hake, 1950; Fabry et al., 1993; Gagné, 1988). To explore the potential influence of presentation level further, Pearson-Product correlations were computed using the difference in presentation level as the independent variable and the change in temporal intra-speech masking (unfiltered vs. spectrally-shaped) as the dependent variable. Significant positive correlations (greater vowel masker level associated with greater intra-speech masking) would support the idea that presentation level may have significantly influenced the results. The correlation coefficients were $-.37$, $-.47$, and $-.16$ for /t/, /p/, /k/, respectively. These were not significant at the .05 level and were in the opposite direction than would be predicted if masker presentation level was influencing the magnitude of intra-speech masking. Thus, there is no evidence that differences in presentation level had a substantial influence on the results.

Results for the unfiltered speech were characterized by substantial inter-subject differences, particularly for listeners with steeply sloping hearing losses. Hearing loss slope accounted for a modest proportion of the variance, but other factors (pure-tone average thresholds and listener age) did not contribute significantly. In addition, there were no significant Pearson Product-Moment correlations between the magnitude of temporal intra-speech masking and listener age or pure-tone thresholds ($p > .01$). The involvement of hearing loss slope as a contributing factor is consistent with the idea that upward spread of masking from the lower frequency vowel peaks produces more masking for more steeply sloping losses because of better hearing sensitivity in the regions of the vowel peaks. The absence of age effects may be due, in part to the relatively restricted age range of the listeners with hearing loss. Other factors contributing to inter-subject differences may include inter-subject variability in frequency selectivity and recovery from forward masking resulting from differences in cochlear pathology.

Because this study addressed only consonant burst detection, the findings cannot be directly generalized to temporal intra-speech masking effects on recognition. Nevertheless, given that detection is required for recognition, the excess temporal intra-speech masking of consonant bursts by adjacent vowels would be expected to disrupt the use of consonant burst cues needed for perception of some speech contrasts, such as consonant place. More work is needed, however, to determine if the release from masking produced by spectral-shaping results in parallel improvements in consonant discrimination.

The findings of this study have several implications for sensory aid processing. First, failure to appropriately shape the frequency response of hearing aids or assistive listening devices for listeners with steeply sloping hearing loss may result in excess temporal intra-speech masking. Cell phones and telephone amplifiers (if used without hearing aids), for example, may not provide appropriate high frequency emphasis without over-amplifying low frequencies. Secondly, even with appropriate spectral shaping, additional signal processing may be needed for individuals who continue to experience excess intra-speech masking. Fast-acting multi-channel compression may be beneficial in these instances.

SUMMARY AND CONCLUSIONS

Temporal masking of consonant bursts by adjacent vowels was measured in listeners with normal hearing and gradually- and steeply sloping sensorineural hearing loss. The conclusions of the study are as follows:

1. Listeners with steeply sloping hearing loss have significantly more temporal intra-speech masking of consonant bursts by adjacent vowels than listeners with flat/gradually sloping hearing loss or normal hearing.

2. Spectral shaping designed to mimic the frequency response of a hearing aid reduces the magnitude of temporal intra-speech masking of /k/ and /t/ for listeners with steeply sloping hearing losses.
3. Though excess masking in listeners with steeply sloping hearing loss is generally reduced by frequency shaping, some excess temporal intra-speech masking remains for some listeners.

References

- Byrne C, Cotton S. Evaluation of the National Acoustic Laboratories' new hearing aid selection procedure. *Journal of Speech and Hearing Research* 1988;31:178–186. [PubMed: 3398491]
- Byrne D, Dillon H. The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear and Hearing* 1986;7:257–265. [PubMed: 3743918]
- Cassidy RL, Harrington J. The place of articulation distinction in voiced stops: Evidence from burst spectra and formant transition. *Phonetica* 1995;52:263–284.
- Cazals Y. Occlusive silence duration of voiceless intervocalic plosives and voicing perception by normal and hearing-impaired subjects. *Ear & Hearing* 1994;15:404–408. [PubMed: 7813827]
- Cazals Y, Palis L. Effect of silence duration in intervocalic velar plosive on voicing perception for normal and hearing-impaired subjects. *Journal of the Acoustical Society of America* 1991;89(6):2916–2921. [PubMed: 1918632]
- Cook JA, Bacon SP, Sammeth CA. Effect of low-frequency gain reduction on speech recognition and its relation to upward spread of masking. *Journal of Speech, Language, & Hearing Research* 1997;40(2):410–422.
- Danaher EM, Osberger MJ, Pickett JM. Some masking effects produced by low-frequency vowel formants in persons with sensorineural hearing loss. *Journal of Speech and Hearing Research* 1973;16:439–451. [PubMed: 4744962]
- Danaher EM, Pickett JM. Some masking effects produced by low-frequency vowel formants in persons with sensorineural hearing loss. *Journal of Speech & Hearing Research* 1975;18(2):261–271.
- Doherty KA, Lutfi RA. Spectral weights for overall level discrimination in listeners with sensorineural hearing loss. *Journal of the Acoustical Society of America* 1996;99(2):1053–1058. [PubMed: 8609288]
- Dorman MF, Lindholm JM, Hannley MT. Influence of the first formant on the recognition of voiced stop consonants by hearing-impaired listeners. *Journal of Speech & Hearing Research* 1985a;28(3):377–380. [PubMed: 2931555]
- Dorman MF, Marton K, Hanley MT, Lindholm JM. Phonetic identification by elderly normal and hearing-impaired listeners. *Journal of the Acoustical Society of America* 1985b;77:664–670. [PubMed: 3973237]
- Egan JP, Hake HW. On the masking pattern of a simple auditory stimulus. *Journal of the Acoustical Society of America* 1950;22:622–630.
- Fabry DA, Leek MR, Walden BE, Cord M. Do adaptive frequency response (AFR) hearing aids reduce 'upward spread' of masking. *Journal of Rehabilitation Research and Development* 1993;30:318–325. [PubMed: 8126656]
- Gagné J-P. Excess masking among listeners with a sensorineural hearing loss. *Journal of the Acoustical Society of America* 1988;83(6):2311–2321. [PubMed: 3411023]
- Gordon-Salant S. Effects of acoustic modification on consonant recognition by elderly hearing-impaired subjects. *Journal of the Acoustical Society of America* 1987;81:1199–1202. [PubMed: 3571732]
- Gehr SE, Sommers MS. Age differences in backward masking. *Journal of the Acoustical Society of America* 1999;106(5):2793–2799. [PubMed: 10573894]
- Hanley M, Dorman MF. Susceptibility to intraspeech spread of masking in listeners with sensorineural hearing loss. *The Journal of the Acoustical Society of America* 1983;74(1):40–51. [PubMed: 6886197]
- Hawkins DB, Walden BE, Prosek RA. Description and validation of an LDL procedure designed to select SSPL90. *Ear and Hearing* 1987;8:162–169. [PubMed: 3609512]

- Levitt H. Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America* 1971;49:467–477. [PubMed: 5541744]
- Lisker L. Closure duration and the intervocalic voiced distinction in English. *Language* 1957;33:42–49.
- Pickett JM. Low-frequency noise and methods for calculating speech intelligibility. *Journal of the Acoustical Society of America* 1959;31:1259–1263.
- Pickett, JM.; Danaher, EM. On discrimination of formant transitions by persons with severe sensorineural hearing loss. In: Fant, G.; Tatham, MAA., editors. *Auditory analysis and perception of speech*. Academic; New York: 1975. p. 275-292.
- Raz I, Noffsinger D. Identification of synthetic, voiced stop-consonants by hearing-impaired listeners. *Audiology* 1985;24(6):437–448. [PubMed: 4084116]
- Revoile S, Pickett JM, Wilson MP. Masking of noise bursts by an adjacent vowel for hearing-impaired listeners. *Journal of Speech & Hearing Research* 1981;24(4):576–579. [PubMed: 7329053]
- Sammeth CA, Birman M, Hecox KE. Variability of most comfortable and uncomfortable loudness levels to speech stimuli in the hearing impaired. *Ear & Hearing* 1989;10:94–100. [PubMed: 2707507]
- Stelmachowicz PG, Kopun J, Mace A, Lewis DE, Nitttrouer S. The perception of amplified speech by listeners with hearing loss: Acoustic correlates. *Journal of the Acoustical Society of America* 1995;98:1388–1399. [PubMed: 7560508]
- Summers V, Leek MR. Intraspeech spread of masking in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America* 1997;101(5 Pt 1):2866–2876. [PubMed: 9165738]
- Tucker-Davis. SigGen Software (Version 1.0). Tucker-Davis Technologies; Gainesville: 1997.
- Van Tasell DJ. Perception of second-formant transitions by hearing-impaired persons. *Ear & Hearing* 1980;1:130–136. [PubMed: 7390071]

Figure 1. Average pure-tone thresholds for participants with flat/gradually sloping hearing loss configurations and steeply sloping configurations. The error bars denote ± 1 standard deviation.

Figure 2.

Spectra for the unfiltered and filtered vowel and consonants: vowel /a/ (broken lines) and the three consonant bursts (solid lines: /k/ (top row), /p/ (middle row), and /t/ (bottom)).

Figure 3.

Mean (bars) and individual (symbols) temporal intra-speech masking for two groups of listeners with hearing loss for the unfiltered stimuli. Data points for each subject are in the same position (left to right) in each panel (e.g. SS01 is on the far left in each panel). The mean and upper limits of the 95% confidence intervals for normal-hearing listeners are indicated by the solid and dashed horizontal lines, respectively.

Figure 4. Temporal speech masking of unfiltered speech stimuli for individual listeners with hearing loss plotted as a function of hearing loss slope.

Figure 5.

Mean temporal intra-speech masking for unfiltered and spectrally-shaped speech for listeners with flat/gradually sloping (F/GS) and steeply sloping losses (SS). Error bars denote +1 standard error. The mean unfiltered temporal intra-speech masking value and upper limit of the 95% confidence interval for normal-hearing listeners are indicated by the solid and dotted lines, respectively.

Figure 6.

Mean consonant burst intensity change with spectral shaping (top panel) and mean consonant-to-vowel ratios for the unfiltered and spectrally-shaped conditions (bottom panel) for listeners with the flat/gradually sloping and steeply sloping hearing loss.

Results of a forward step-wise regression of unfiltered intra-speech masking for listeners with hearing loss. The independent variables were age, pure-tone average, and hearing loss slope (dB / octave between 500 & 4000 Hz). Only variables with F-ratios greater than 1.0 were included in the model

Table 1

	Multiple R		R-Squared		F-ratio	p-level	Step
	Multiple R	R-Squared	change	R-Squared			
/t/ burst slope	0.550	0.303	0.303	0.303	5.217	0.041	1
age	0.611	0.373	0.070	0.373	1.236	0.290	2
/p/ burst slope	0.523	0.273	0.273	0.273	4.515	0.055	1
/k/ burst slope	0.494	0.244	0.244	0.244	3.876	0.073	1