NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Age effects in temporal envelope processing: speech unmasking and auditory steady state responses

John H. Grose, Ph.D., **Sara K. Mamo, Au.D.**, and **Joseph W. Hall III, Ph.D.**
Dept. Otolaryngology – Head & Neck Surgery, University of North Carolina at Chapel Hill

Introduction

Auditory temporal processing abilities tend to decline with advanced age, and independently of hearing loss (for review, see Gordon-Salant, 2006). Although this holds for many aspects of temporal processing, consensus is lacking with regard to the processing of on-going amplitude modulations, or temporal envelopes. Temporal envelope refers to the slower amplitude fluctuation that is carried by the more rapid pressure oscillations, or fine structure, of a sound. The purpose of this study is to address the question of envelope processing in older listeners in an attempt to forge consensus on this issue.

Several studies suggest that older listeners do not exhibit deficits in temporal envelope processing. Peters and Hall (1994) found that, for tone detection in a sub-critical band of noise, the advantage conferred by square-wave modulating the noise amplitude was not diminished in older listeners. That is, the modulated-unmodulated threshold difference across a 10 – 50 Hz range of modulation rates exhibited no age effect. The ability to benefit from modulated noise depends in part on the ability to resolve the masker minima; i.e., it requires that the masker fluctuation pattern be followed with fidelity. In similar vein, Takahashi and Bacon (1992), using sinusoidal modulation, found that differences across age groups for both modulation detection (temporal modulation transfer functions [TMTFs]), as well as the benefit to masked speech recognition of modulating the masker at 8 Hz, was more a function of the audiometric hearing status of the listeners than their age per se. Speech recognition benefit in modulated maskers again relies in part on accurate coding of the masker envelope. Other evidence unsupportive of an age effect in temporal envelope processing comes from electrophysiological work. Boettcher et al. (2001) measured the 40-Hz amplitude modulation following response (AMFR) in younger and older listeners and found no differences in response amplitude over a wide range of modulation depths. These studies therefore suggest that, in the absence of peripheral sensory loss, older listeners exhibit the same envelope processing capabilities as younger listeners.

In contrast, several studies have highlighted envelope processing deficits in older listeners. In terms of speech recognition in modulated noise, a number of studies have shown age-related reductions in masker modulation benefit (e.g., Dubno et al., 2003; George et al., 2007; Gifford et al., 2007). In these studies, square-wave modulation rates ranged from 2 – 50 Hz and the speech material ranged from nonsense syllables to sentences. A study of TMTFs using tonal carriers found that older listeners had elevated thresholds relative to younger listeners for modulation frequencies ≥ 40 Hz, except at very high frequencies where performance was presumably based entirely on spectral cues; i.e., side-band resolution (He et al., 2008). Some electrophysiological evidence also suggests envelope processing deficits in older listeners.

Address for correspondence: John H. Grose, Dept. OHNS, UNC-CH, 126 Physicians Office Building, CB#7070, 170 Manning Drive, Chapel Hill, NC 27599-7070, Voice: (919) 966-8926, Fax: (919) 966-7656, jhg@med.unc.edu.

Leigh-Paffenroth and Fowler (2006) measured the auditory steady state response (ASSR) in younger and older listeners using a 'present/absent' response metric based on the phase coherence of the electroencephalographic component at the modulation frequency. They found that younger listeners had more phase-locked responses overall than older listeners for a 500-Hz carrier (but not for a 2000-Hz carrier), and that, across the two carrier frequencies, the older listeners also had less phase-locked responses at the highest rate tested of 90 Hz. A study by Purcell et al. (2004) included both psychophysical and electrophysiological measures. In both domains, a white noise was sinusoidally modulated at a depth of 25% and the modulation frequency was varied. The highest modulation rate at which the modulation could be detected psychophysically was measured, as well as the highest modulation rate that elicited a detectable envelope following response (EFR). [Note that the terms AMFR, EFR and ASSR can be used interchangeably; the term ASSR will be used hereafter.] They found that the younger listeners could detect the presence of 25% modulation to a higher modulation frequency (567 Hz) than the older listeners (264 Hz), and that the highest detectable ASSR modulation frequency was also significantly higher in younger (494 Hz) than older (294 Hz) listeners; the two groups diverged in ASSR amplitude above about 100 Hz. These studies therefore suggest that older listeners do exhibit deficits in envelope processing, whether measured psychophysically or electrophysiologically.

Although the literature contains divergent findings with regard to the effect of age on envelope processing, one trend that emerges from these studies is that age effects are more likely to be present at higher envelope frequencies. For example, the absence of an age effect found psychophysically by Peters and Hall (1994) and electrophysiologically by Boettcher et al. (2001) employed modulation rates ≤ 50 Hz. He et al. (2008) also found no age effect at the lowest modulation frequency tested of 5 Hz. Likewise, Purcell et al. (2004) found no age effect in the amplitude of the ASSR response when the modulation frequency ranged from 30 – 50 Hz. In contrast, the modulation frequency that generated the largest ASSR age effect in the Leigh-Paffenroth and Fowler (2006) study was 90 Hz, and the separation between younger and older listeners found psychophysically and electrophysiologically by Purcell et al. (2004) occurred for modulation rates over 100 Hz. One hypothesis that can be constructed from this pattern of data, therefore, is that advanced age affects measures of temporal envelope processing at high modulation rates but not at low rates. A purpose of this study was to test this hypothesis; specifically, this study tested whether ASSR amplitudes are reduced in older listeners for high modulation rates but not for low modulation rates.

As suggested earlier, it is reasonable to suppose that the benefit to masked speech recognition of modulating the masker depends in part on the fidelity with which the masker envelope is encoded. In consequence, it might be expected that the recognition benefit should diminish as modulation rate is increased above the region of optimum modulation sensitivity. For broadband stimuli, this cutoff frequency is about 50 Hz (e.g., Viemeister, 1979). However, Dubno et al. (2003) found that, for both younger and older listeners, the recognition benefit for speech presented at a moderate level in modulated noise declined once the modulation rate exceeded 25 Hz. It is therefore possible that other factors in addition to temporal envelope coding play a role in this effect. One possibility is that the limiting factor in speech recognition in modulated noise is the relative sparsity of available speech cues which have become constrained in modulated noise due to the limited 'windows' during which the speech can be glimpsed (Buss et al., 2004; Cooke, 2006; Miller and Licklider, 1950). Another way of expressing this reduction in available speech cues is as a reduction in the redundancy of the speech signal. Speech redundancy refers to the multiplicity of coexisting speech cues, including contextual, coarticulatory, and other acoustic cues. Constraining these cues can therefore be viewed as a reduction in speech redundancy. If this reduced speech redundancy affects older listeners more than younger listeners, a detriment in speech recognition benefit will be observed. For example, suppose that the cues for speech recognition are effectively less

redundant for older listeners than younger listeners even in optimal listening conditions. This could arise in a number of ways including degraded coding of individual cues or a reduced ability to combine multiple cues. Nonetheless, as long as sufficient cues for speech recognition remain available, the performance of older listeners will not show any detriment. However, once the listening conditions constrain the cues still further, the older listeners will now be proportionally more affected than younger listeners. In the context of masked speech recognition, it is possible that younger and older listeners exhibit similar performance in a steady masker because sufficient speech cues are continuously available to each group. That is, there is no temporal interruption in the availability of the sufficient cues. However, in the modulated masker where speech recognition is presumably based mainly on the fragments of speech available during the periods of advantageous speech-to-masker ratios in the masker minima, the cue set becomes more constrained by virtue of the temporal interruptions. This results in a performance detriment in the older listeners relative to the younger listeners. The net result is a reduced speech recognition advantage in modulated noise for the older listeners. If this suggestion is reasonable, then the benefit to speech recognition of modulating the masker should be sensitive not only to the rate of masker modulation but also to the relative redundancy of the speech for a given listener. One way to test this would be to assess the magnitude of speech unmasking (i.e., the benefit to speech recognition of modulating the masker) as a function of the inherent redundancy of the speech material. This forms the second purpose of this study – to test the effects of age on speech unmasking as a function of the masker modulation rate and the redundancy of the speech material. The hypothesis is that older listeners will exhibit reduced speech unmasking relative to younger listeners at higher masker modulation rates and especially for less redundant speech. As expanded on below, speech redundancy was manipulated by varying speech rate.

In summary, the purpose of this study was to determine whether temporal envelope processing is degraded in older listeners. Two experiments were undertaken. The first examined speech unmasking in modulated noise as a function of modulation rate and speech redundancy. A goal of this experiment was to ascertain whether performance deficits of older listeners in modulated noise related to poor temporal coding of the modulated masker envelope or to reduced speech redundancy. The second experiment measured ASSR amplitudes at a low and a high modulation rate for two carrier frequencies. The purpose of this experiment was to test the hypothesis that older listeners exhibit reduced amplitudes at high modulation rates but not at low rates. By testing the same listeners across the two experiments, both behavioral and electrophysiological information could be brought to bear on the question of temporal envelope processing in older listeners.

Exp. 1. Speech unmasking as a function of speech rate and masker modulation rate

The purpose of this experiment was to test the hypothesis that older listeners exhibit reduced speech unmasking relative to younger listeners at higher masker modulation rates, especially for speech with reduced redundancy. As noted above, there are several dimensions along which the redundancy of speech can be varied such as contextual integrity (e.g., high-predictability *vs.* low-predictability speech) and acoustic integrity (e.g., filtered *vs.* unfiltered speech). In this study, speech redundancy was varied by manipulating the degree of temporal compression, or speech rate. Time compression of speech can be implemented in a number of ways; e.g., dropping every 'nth' sample of the digital waveform or culling every 'nth' frame of a steady-state segment. The degree of compression is usually defined as the percentage by which the duration of the original speech waveform is reduced. Each compression algorithm comes with associated costs in terms of signal distortion. Because the algorithms involve removal of parts of the speech waveform, they inherently reduce the acoustic redundancy of the speech signal. Several studies have shown that the ability to recognize time-compressed speech is often

reduced in older listeners (e.g., Gordon-Salant and Fitzgibbons, 2004; Jenstad and Souza, 2007; Wingfield et al., 2006). However, whether older listeners have difficulty with time-compressed speech depends on a complex interaction of factors including the compression rate, compression method, and type of speech material. For example, Vaughan and Letowski (1997) found that older listeners did not show substantial deficits for repeating highly predictable sentences until the temporal compression rate was increased above 60%. In contrast, time compression of even 50% can result in performance deficits for low-predictability sentence material (e.g., Gordon-Salant et al., 2007). Thus, the effects of time compression depend on contextual redundancy. The acoustic characteristics of the time-compressed speech can also affect performance. In a study that compared the effects of different compression algorithms, Schneider et al. (2005) found that recognition of time-compressed speech by older listeners was minimally affected if the compression algorithm removed only steady-state segments, preserving key speech features. In that study, speech recognition was measured in a 12-talker babble background that was otherwise unmodulated. A similar compression technique was chosen here (see below) to minimize signal distortion. Finally, although it is intuitive that deficits in perceiving rapid speech might also reflect a general slowing of processing speed in older adults, Jenstad and Souza (2007) have shown that rapid speech deficits are associated primarily with the loss of acoustic redundancy in the speech signal rather than with a loss of processing speed in the listener. In summary, the purpose of this experiment was to assess speech unmasking as a function of listener age to gauge the fidelity with which the temporal envelope is processed. Speech unmasking was measured for two masker modulation rates and, in addition, the redundancy of the speech was manipulated using time compression.

Method

Observers

The observers were young (n = 10) and older (n = 10) adults with relatively normal hearing. The young observers ranged in age from 21 to 29 yrs (mean = 25.0 yrs) and had audiometric thresholds ≤ 20 dB HL across the octave frequencies 250 – 8000 Hz. The older observers ranged in age from 63 to 75 yrs (mean = 68.7 yrs) and had audiometric thresholds ≤ 20 dB HL across the octave frequencies 250 – 4000 Hz, with the exception of two observers who had thresholds of 25 dB HL at 4000 Hz. Most older observers had some threshold elevation at higher frequencies. Table 1 lists the mean audiometric thresholds, and standard deviations, across frequency for the two groups.[1]

Stimuli

The target speech consisted of the revised Harvard list of phonetically balanced sentences, known as the IEEE sentences [72 lists of 10 sentences each] (IEEE, 1969). The master corpus was a pre-recorded version of the speech material (Loizou, 2007) which was transferred to a computer hard drive for off-line processing. After down-sampling to a rate of 24,414 Hz, one copy of the speech material was left unmodified while a second copy was submitted to a time compression algorithm. Time compression was applied using a Waveform Similarity OverLap Add (WSOLA) algorithm [SoundTouch©, Espoo, Finland]. This processing compresses the time waveform without altering its pitch by removing brief segments of the waveform that are replicated in time (i.e., redundant) such as periodicity segments. The algorithm was set to increase the tempo of the signal by 50%, which resulted in a reduction of the duration of each speech waveform to 67% of its original length. The 33% time-compressed speech will be referred to hereafter as 'rapid speech.'

---

[1]Although the older listeners generally had slightly poorer audiometric thresholds relative to the younger listeners, none of the experimental data reported in this study showed any association with observers' audiometric sensitivity.

The target speech was masked by a continuous broadband noise shaped to have the same long-term spectrum as the sentences. The masker was either presented in its steady mode (i.e., "unmodulated"), where no extraneous modulation was imposed, or it was square-wave modulated at a rate of 16 Hz or 32 Hz. Masker level was defined as the RMS of the unmodulated masker in dB SPL. No additional level adjustment was applied to the modulated masker.

Procedure

Masked speech thresholds were measured in a computerized adaptive procedure that varied the masker level while keeping the target speech level fixed. In each trial of an adaptive threshold track, a target sentence was selected from the appropriate corpus (normal or rapid speech) and presented to the observer at a fixed level of 65 dB SPL. For each observer, the initial sentence in the experiment was selected randomly from within the appropriate corpus and subsequent sentences progressed sequentially from that point across trials and across tracks. No observer received the same sentence twice, and no two observers received completely overlapping sequences of sentences. After each presentation, the observer repeated the perceived sentence aloud, and the experimenter (outside the booth monitoring the response over headphones via a talk-back loop) scored it as correct/incorrect, where missing any of the 5 keywords/sentence resulted in a score of incorrect. This score was entered into the adaptive program via a computer mouse click and the masker level for the next trial was adjusted based on the stepping rule. The masker level was increased by 2 dB after a correct response, and decreased in level by 2 dB after an incorrect response, resulting in a convergence on the masker level at which the speech was recognized at the 50% correct point. The adaptive threshold track was terminated after 6 reversals in masker level and the estimate of threshold for that track was taken as the mean of the final 4 reversal levels. Note that in this method of fixed signal level and variable masker level, higher thresholds denote better performance since they indicate that the observer is able to tolerate greater masker energy while still correctly perceiving the target sentence. Masked thresholds for both normal and rapid speech were measured for three masker modulation rates: 0 Hz [steady masker], 16 Hz, and 32 Hz. For each condition, three estimates of threshold were measured, with a fourth collected if the range of the first three was > 3 dB. The final threshold for that condition was taken as the mean of all estimates collected. The conditions were blocked for an observer but the order of conditions across observers was random. No specific training was incorporated. Stimuli were presented monaurally under Sennheiser HD580 headphones in a sound-attenuating booth. The left ear was selected as the test ear unless the audiometric sensitivity in the right ear better met the inclusion criteria; five of the older observers were tested in the right ear.

Results and Discussion

The masker levels at threshold are plotted as a function of masker modulation rate in Fig. 1. Panels A and B show individual and mean data for normal and rapid speech, respectively, for both younger (circles) and older (squares) listeners. The mean data and associated standard deviations are re-plotted for all conditions in panel C for both age groups, with filled symbols indicating normal speech and open symbols indicating rapid speech. The modulation masking release (speech unmasking) is plotted as a function of modulation rate in Fig. 2. This portrays the difference in thresholds between the steady masker (0-Hz modulation) and the masker modulated at either 16 Hz or 32 Hz. Again, individual and mean data are shown in Panels A and B for normal and rapid speech, respectively, with Panel C re-plotting all mean data. The same labeling conventions are used as in Fig. 1. The key observations from these data can be summarized as follows:

1) Speech unmasking as a function of age

The benefit to speech recognition of modulating the masker, as a function of listener age, was assessed by comparing the derived masking release for normal speech across the two age groups for both masker modulation rates (see Fig. 2). The masking release magnitudes were submitted to an ANOVA with one between-subjects factor (Age: Younger, Older) and one within-subjects factor (Modulation rate: 16 Hz, 32 Hz). The analysis indicated a significant effect of age ($F_{1,18} = 11.02$; $p < 0.01$) and a significant effect of modulation rate ($F_{1,18} = 5.91$; $p = 0.026$); the interaction between these two factors was not significant ($F_{1,18} = 1.09$; $p = 0.31$). This result indicates that older listeners do not benefit as greatly as younger listeners when the masker is modulated (i.e., they show less speech unmasking), and that both age groups exhibit less benefit for a 32-Hz modulation rate than for a 16-Hz modulation rate. It is evident from Fig. 2 that one older listener exhibited particularly low speech unmasking; however, repeating the analysis with this observer excluded did not change the pattern of effects. Reduced speech unmasking in older listeners has been found in several studies (e.g., Dubno et al., 2003;George et al., 2007;Gifford et al., 2007), and the reduction in magnitude of speech unmasking with increasing modulation rate, irrespective of listener age, parallels the finding of Dubno et al. (2003).

2) Effect of speech rate as a function of age

The effect of speech rate on masked speech recognition, as a function of age, was assessed by comparing sentence recognition performance across the two age groups for normal and rapid speech presented in the steady masker (see Fig. 1). The masker levels at threshold were submitted to an ANOVA with one between-subjects factor (Age: Younger, Older) and one within-subjects factor (Speech type: Normal, Rapid). The analysis indicated a significant effect of age ($F_{1,18} = 9.39$; $p < 0.01$), a significant effect of speech type ($F_{1,18} = 165.93$; $p < 0.001$), and a significant interaction between the two factors ($F_{1,18} = 19.78$; $p < 0.001$). Post-hoc analysis of the interaction term indicated that the two age groups did not differ for the normal speech ($F_{1,18} = 0.08$; $p = 0.78$) but did for the rapid speech ($F_{1,18} = 16.23$; $p = 0.001$). This result indicates that increased age does not affect the perception of normal speech presented in a steady masker, but that older listeners are poorer at perceiving rapid speech under the same masking condition. Reduced recognition of rapid speech by older listeners has been found in several studies (e.g., Gordon-Salant and Fitzgibbons, 2004;Jenstad and Souza, 2007;Wingfield et al., 2006), although this depends on the manner of temporal compression (Schneider et al., 2005).

3) Age effects as a function of speech rate and modulation rate

The analyses above highlight several key points: (a) for normal speech presented in a steady masker, older listeners show no detriment in speech recognition performance relative to younger listeners; (b) when the masker for the normal speech is modulated, the older listeners show less speech unmasking relative to younger listeners; and (c) when the masker remains steady but the speech rate is increased, the older listeners show a greater performance deficit than the younger listeners. These points generally confirm previous findings, and the main question of interest in this experiment was how these two dimensions (speech rate and masker modulation rate) interact with each other as a function of listener age. To test this, the magnitudes of masking release were compared for the various combinations of speech rate and modulation rate for the two age groups. The ANOVA had one between-subjects factor (Age: Younger, Older) and two within-subjects factors (Speech rate: Normal, Rapid; and Modulation rate: 16 Hz, 32 Hz). The analysis revealed no main effect of age ($F_{1,18} = 3.17$; $p = 0.092$) but a significant interaction of age and speech rate ($F_{1,18} = 12.95$; $p = 0.002$). The main effect of speech rate itself was significant ($F_{1,18} = 80.5$; $p < 0.001$) and, in addition to the interaction just noted, the interaction between speech rate and modulation rate was also significant

($F_{1,18} = 5.45$; p = 0.031). The main effect of modulation rate was not significant ($F_{1,18} = 1.09$; p = 0.31), nor was its interaction with age group ($F_{1,18} = 0.013$; p = 0.91). Finally, the 3-way interaction of age*speech type*modulation rate was also not significant ($F_{1,18} = 2.9$; p = 0.11). Post-hoc analysis of the interaction between age and speech rate indicated that the older listeners exhibited less masking release than the younger listeners for normal speech ($F_{1,18} = 11.02$; p < 0.01), as already noted above, but that the two age groups did not differ significantly in the magnitude of masking release for the rapid speech ($F_{1,18} = 0.085$; p = 0.77). Analysis of the interaction between speech rate and modulation rate indicated that, for the normal speech, there was significantly more masking release observed at 16-Hz modulation than at 32-Hz modulation ($F_{1,18} = 5.91$; p = 0.03), again as described earlier, but that there was no effect of modulation rate for the rapid speech ($F_{1,18} = 0.38$; p = 0.55). In summary, these results indicate that for rapid speech there is no age effect for masking release and there is no difference in the magnitude of masking release for the two modulation rates. Effects of listener age and masker modulation rate on the magnitude of masking release are observed only for normal speech.

Exp. 1 confirms that, for normal speech, older listeners exhibit less speech unmasking than younger listeners. A key consideration is whether this demonstrates a deficit in temporal envelope coding (of the masker) or whether it reflects some other factor perhaps related to the redundancy of the available speech. We contend that the data pattern is consistent with an interpretation in terms of reduced speech redundancy rather than deficits in envelope coding. That is, the failure of older listeners to benefit from a modulated masker to the same extent as younger listeners for normal speech is likely because the effective speech cues during the masker minima are more restricted for the older listeners, not because the minima are less resolved. The logic of this interpretation is based on the premise that younger listeners with normal hearing exhibit optimum modulation processing of low frequency envelopes (as measured, for example, by TMTFs). For these listeners, it can be assumed that envelope processing is equivalent for the two modulation rates of 16 Hz and 32 Hz. This premise leads to two conclusions regarding the speech performance of the younger listeners: (1) for normal speech, the reduced benefit of modulating the masker at 32 Hz *vs.* 16 Hz must be due to a reduced ability to reconstruct the speech signal from the available 'glimpses' during the shorter and more frequent masker minima; (2) the reduced benefit of modulating the masker for rapid speech relative to normal speech must be due to poorer residual speech cues existing during the masker minima in the case of rapid speech. These two conclusions guide the interpretation of the performance of older listeners in the same conditions. First, for normal speech the older listeners also exhibit a reduction in the benefit of modulating the masker at 32 Hz vs. 16 that parallels that of the younger listeners. This parallel effect of masker modulation rate across the two age groups for normal speech suggests that the decline observed in the older listeners occurs for the same reason as it does for the younger listeners; i.e., a reduced ability to reconstruct the speech signal from the available 'glimpses' during the shorter and more frequent masker minima, and not to a deficit in coding the masker envelope, per se. Second, for rapid speech relative to normal speech, the older listeners actually show *less* reduction in masker modulation benefit compared to younger listeners; i.e., collapsed across modulation rates the reduction is 7 dB for the younger listeners but only 3 dB for the older listeners. This suggests that the magnitude of the masker modulation benefit for the older listeners is driven more by the inherent quality of speech cues available than by the resolution of the masker envelope itself. In summary, the results of exp. 1 suggest that the reduced speech unmasking seen in older listeners for relatively slow modulation rates is not due to deficits in envelope processing but rather is associated with the more constrained redundancy of the speech material.

Exp. 2. ASSR amplitude as a function of carrier and modulation frequencies

The purpose of this experiment was to undertake an electrophysiological test of the hypothesis that advanced age affects measures of temporal envelope processing at high modulation rates

but not at low rates. The electrophysiological test consisted of the measurement of ASSR amplitude at a low and a high modulation rate. The low modulation rate was 32 Hz, which corresponds to one of the masker modulation rates used in exp. 1 and is also in the general vicinity of the rate used in a previous ASSR study where no age effect was observed (Boettcher et al., 2001). The high modulation rate was 128 Hz, which is within the range of frequencies where two previous ASSR studies had observed age effects (Leigh-Paffenroth and Fowler, 2006; Purcell et al., 2004).

Method

Observers

The same 20 observers from exp. 1 participated in exp. 2.

Stimuli

ASSRs were elicited with sinusoidally amplitude modulated (SAM) tones having carrier frequencies (fc) of 500 Hz or 2000 Hz and modulator frequencies (fm) of 32 Hz or 128 Hz. Modulation depth was 100%. Stimuli were presented monaurally at a level of 75 dB HL. For each observer, the test ear was the same as that used in exp. 1

Procedure

Observers relaxed in a reclining chair in a sound-attenuating booth and watched a silent movie of their choice with captioning. Single channel recordings were made ipsilateral to the test ear using the BioLogic MASTER™ research system [Mundelein, IL, USA]. The non-inverting electrode was placed at Fz, the inverting electrode on the ipsilateral earlobe, and the ground electrode at Fpz. Electrode impedances were maintained at ≤ 3 kΩ. The MASTER system band-pass filters the recorded EEG from 1 – 300 Hz, and defaults to collect a maximum of 32 sweeps of 16 epochs/sweep, where each epoch consists of a 1024-point time segment sampled at a rate of 1000 Hz (John and Picton, 2000). An artifact rejection criterion of 40 μV was employed. Following every 16 epochs, the display of the EEG spectrum in the region of the modulation frequency is refreshed, along with an update on the probability that the magnitude of the spectral component at fm is drawn from the same distribution as the surrounding noise (F statistic corresponding to the signal-to-noise ratio). A run was terminated prior to the default stopping criterion if the F ratio remained significant for several consecutive sweeps *and* the background noise level was less than about 20 nV. If the component amplitude at fm was not significantly different from the noise floor after 32 sweeps of 16 epochs/sweep, the response was considered absent. For each of the four combinations of fc and fm, two replications of ASSR amplitude were collected and averaged.

Results and Discussion

The results of exp. 2 are shown in Fig. 3 where ASSR amplitude is plotted for each combination of fc and fm. Mean data, with associated standard deviations, are shown as large symbols, with individual data offset to either side for younger (circles) and older (squares) listeners. Note that for four cases among the older listeners, the EEG component at $fm = 128$ Hz was not significantly different from the surrounding EEG noise. In each of these four cases, the failure to detect a significant ASSR was not due to an elevated noise floor. The amplitudes recorded for these cases are identified as inverted triangles. In order to determine whether there was an effect of age as a function of either fc or fm, the ASSR amplitudes were submitted to an ANOVA with one between-subjects factor (Age: Younger, Older) and two within-subjects factors (Carrier frequency: 500 Hz, 2000 Hz; Modulation frequency: 32 Hz, 128 Hz). Two points should be noted: (1) The statistical analyses were performed on the log transforms of the data. Although the standard errors of skewness of the ASSR amplitude distributions across

observers for each modulation rate were always less than 0.38 on both linear and logarithmic scales, the skewness was always smaller for the logarithmic transform and therefore this scale was selected for analysis. (2) For the 4 instances where the amplitude of the spectral component at $fm = 128$ Hz failed to differ significantly from the surrounding noise floor, the measured amplitude of the component at this frequency was entered into the data set. The results of the analysis revealed no overall effect of age group ($F_{1,18} = 1.56$; $p = 0.23$) but a significant interaction of age and fm ($F_{1,18} = 9.62$; $p < 0.01$). The main effects of fc and fm were significant ($F_{1,18} = 21.1$ and 134.27, respectively; $p < 0.01$), as was the interaction between these two factors ($F_{1,18} = 15.0$; $p < 0.01$). The interaction between age group and fc was not significant ($F_{1,18} = 3.57$; $p = 0.08$) and the 3-way interaction between age*fc*fm also failed to reach significance ($F_{1,18} = 4.29$; $p = 0.053$). Post-hoc analysis of the interaction between age and fm indicated that the ASSR amplitudes of the two age groups did not differ for $fm = 32$ Hz ($F_{1,18} = 0.71$; $p = 0.41$) but, for $fm = 128$ Hz, the amplitudes of the older listeners were significantly reduced ($F_{1,18} = 8.23$; $p = 0.01$). Analysis of the interaction between fc and fm indicated that the amplitude difference between $fm = 32$ Hz and $fm = 128$ Hz was highly significant at both carrier frequencies ($F_{1,18} = 129.09$ & 96.27, respectively; $p < 0.001$). In summary, this pattern of results shows that the ASSR amplitudes of older listeners are decreased relative to young listeners for high modulation rates ($fm = 128$ Hz) but not for low modulation rates ($fm = 32$ Hz) at both carrier frequencies tested.

Two other subsidiary analyses are of interest. The first subsidiary analysis dealt with the comparative ASSR noise floor amplitudes across the two groups (collapsed across fc). This analysis indicated no main effect of age group ($F_{1,18} = 1.41$; $p = 0.25$) but a significant effect of fm ($F_{1,18} = 50.69$; $p < 0.01$) and a significant interaction between age and fm ($F_{1,18} = 10.0$; $p = 0.005$). Post-hoc analysis of this interaction indicated that, for both age groups, the noise floor was lower for the 128-Hz rate than for the 32-Hz rate ($F_{1,18} = 7.85$ & 52.81 for the younger and older observers, respectively; $p \leq 0.01$). However, of greater interest was the finding that the two age groups did not differ in their noise floor amplitudes at the 32-Hz rate ($F_{1,18} = 0.08$; $p = 0.79$) but did at the 128-Hz rate ($F_{1,18} = 6.46$; $p = 0.02$) – with the older group having the *lower* noise floor. This latter finding prompted a second subsidiary analysis addressing the number of sweeps collected at each fm. Independent t-tests on the number of sweeps collected at each fm for the two age groups revealed a significantly greater number of sweeps collected for the older listeners at 128 Hz ($t_{18} = 3.15$; $p = 0.006$) but not at 32 Hz ($t_{18} = 0.65$; $p = 0.53$). Therefore, the finding of reduced ASSR amplitudes at 128 Hz in the older listeners is accompanied by reduced noise floor amplitudes which may, in turn, reflect an increase in the required number of collection sweeps relative to younger listeners to reach termination criterion.

The results of this experiment are compatible with both the Boettcher et al. (2001) study, which failed to find an age effect, and the Leigh-Paffenroth and Fowler (2006) and Purcell et al. (2004) studies, which did find an age effect. Specifically, the Boettcher et al. (2001) study tested only at a relatively low rate of 40 Hz, not dissimilar from the 32-Hz rate which showed no age effect here, whereas the other two studies also tested at higher rates, with the Purcell et al. (2004) study encompassing the 128-Hz rate which showed an age effect here. Thus it appears that there is general consensus across electrophysiological studies that advanced age can result in deficits in temporal envelope processing, but only at high rates.

General Discussion and Conclusion

The conclusion from electrophysiological data that temporal envelope processing deficits are present in advanced age, but only at high rates, is generally compatible with psychophysical findings. The Purcell et al. (2004) study showed age differences in the highest modulation rate that could be behaviorally detected and He et al. (2008) also observed age-related deficits in

high-frequency envelope processing as measured by tonal TMTFs. The general compatibility between electrophysiological and psychophysical findings would suggest an association between these two measures. In the present study the two experiments included the common envelope frequency of 32 Hz (square-wave modulation of the speech masker in exp. 1 and sinusoidal modulation of the ASSR stimulus in exp. 2). However, an examination of the correlations among the various measures associated with the 32-Hz modulation rate across the two experiments showed no significant relationships. These measures included: (1) the masker level at speech reception threshold; (2) the derived speech benefit in modulated noise; and (3) the amplitude of the ASSR components at 32 Hz expressed in both linear and logarithmic units. Indeed, examination of the complete matrix of correlations among variables in exps. 1 and 2 revealed only one significant correlation – that between the amplitude of the 32-Hz ASSR at 500 Hz and the masker level at speech reception threshold for the unmodulated masker ($r = 0.50$, $p = 0.026$). The interpretation of this correlation is not immediately clear.

In summary, the purpose of this study was to determine whether temporal envelope processing deficits exist in older listeners. The first experiment examined temporal envelope processing as it relates to speech unmasking in modulated noise. The experiment tested the hypothesis that older listeners exhibit reduced speech unmasking relative to younger listeners at higher masker modulation rates and especially for less redundant speech. This hypothesis was only partially supported. Whereas the older listeners did show reduced unmasking for normal speech at the higher rate of 32 Hz, they showed the same reduction at the slower rate of 16 Hz. Moreover, in contrast to the hypothesis, older listeners did not show an exacerbated deficit in unmasking for the less redundant speech. The results were interpreted as indicating that deficits in speech unmasking in modulated noise in older listeners are due, not to deficits in envelope processing of the masker per se, but rather to a reduced redundancy of the speech material. The second experiment was designed to test the hypothesis that advanced age affects temporal envelope processing at high modulation rates but not at low rates. The results supported the hypothesis, showing reduced ASSR amplitudes in older listeners relative to young listeners for a high modulation rate but not for a low modulation rate irrespective of carrier frequency. The general conclusion is that deficits in temporal envelope processing are evident in advanced age, but only for relatively high envelope frequencies.

Acknowledgments

References

Boettcher FA, Poth EA, Mills JH, et al. The amplitude-modulation following response in young and aged human subjects. Hear Res 2001;153:32–42. [PubMed: 11223295]

Buss E, Hall JW 3rd, Grose JH. Spectral integration of synchronous and asynchronous cues to consonant identification. J Acoust Soc Am 2004;115:2278–2285. [PubMed: 15139639]

Cooke M. A glimpsing model of speech perception in noise. J Acoust Soc Am 2006;119:1562–1573. [PubMed: 16583901]

Dubno JR, Horwitz AR, Ahlstrom JB. Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. J Acoust Soc Am 2003;113:2084–2094. [PubMed: 12703719]

George EL, Zekveld AA, Kramer SE, et al. Auditory and nonauditory factors affecting speech reception in noise by older listeners. J Acoust Soc Am 2007;121:2362–2375. [PubMed: 17471748]

Gifford RH, Bacon SP, Williams EJ. An examination of speech recognition in a modulated background and of forward masking in younger and older listeners. J Speech Lang Hear Res 2007;50:857–864. [PubMed: 17675591]

Gordon-Salant S. Speech perception and auditory temporal processing performance by older listeners: Implications for real-world communication. Seminars in Hearing 2006;27:264–268.

Gordon-Salant S, Fitzgibbons PJ. Effects of stimulus and noise rate variability on speech perception by younger and older adults. J Acoust Soc Am 2004;115:1808–1817. [PubMed: 15101658]

Gordon-Salant S, Fitzgibbons PJ, Friedman SA. Recognition of time-compressed and natural speech with selective temporal enhancements by young and elderly listeners. J Speech Lang Hear Res 2007;50:1181–1193. [PubMed: 17905904]

He N, Mills JH, Ahlstrom JB, et al. Age-related differences in the temporal modulation transfer function with pure-tone carriers. J Acoust Soc Am 2008;124:3841–3849. [PubMed: 19206810]

IEEE. IEEE recommended practice for speech quality measurements [Subjective Measurements Subcommittee]. IEEE Trans Audio Electroacoust, AU-17 1969:225–246.

Jenstad LM, Souza PE. Temporal envelope changes of compression and speech rate: combined effects on recognition for older adults. J Speech Lang Hear Res 2007;50:1123–1138. [PubMed: 17905900]

John MS, Picton PE. MASTER: a Windows program for recording multiple auditory steady-state responses. Computer Meth Prog Biomed 2000;61:125–150.

Leigh-Paffenroth ED, Fowler CG. Amplitude-modulated auditory steady-state responses in younger and older listeners. J Am Acad Audiol 2006;17:582–597. [PubMed: 16999253]

Loizou, PC. Speech Enhancement: Theory and Practice. Boca Raton, FL: CRC Press; 2007.

Miller GA, Licklider JCR. The intelligibility of interrupted speech. J Acoust Soc Am 1950;22:167–173.

Peters RW, Hall JW 3rd. Comodulation masking release for elderly listeners with relatively normal audiograms. J Acoust Soc Am 1994;96:2674–2682. [PubMed: 7983273]

Purcell DW, John SM, Schneider BA, et al. Human temporal auditory acuity as assessed by envelope following responses. J Acoust Soc Am 2004;116:3581–3593. [PubMed: 15658709]

Schneider BA, Daneman M, Murphy DR. Speech comprehension difficulties in older adults: cognitive slowing or age-related changes in hearing? Psychol Aging 2005;20:261–271. [PubMed: 16029090]

Takahashi GA, Bacon SP. Modulation detection, modulation masking, and speech understanding in noise in the elderly. J Speech Hear Res 1992;35:1410–1421. [PubMed: 1494284]

Vaughan NE, Letowski T. Effects of age, speech rate, and type of test on temporal auditory processing. J Speech Lang Hear Res 1997;40:1192–1200. [PubMed: 9328889]

Viemeister NF. Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am 1979;66:1364–1380. [PubMed: 500975]

Wingfield A, McCoy SL, Peelle JE, et al. Effects of adult aging and hearing loss on comprehension of rapid speech varying in syntactic complexity. J Am Acad Audiol 2006;17:487–497. [PubMed: 16927513]
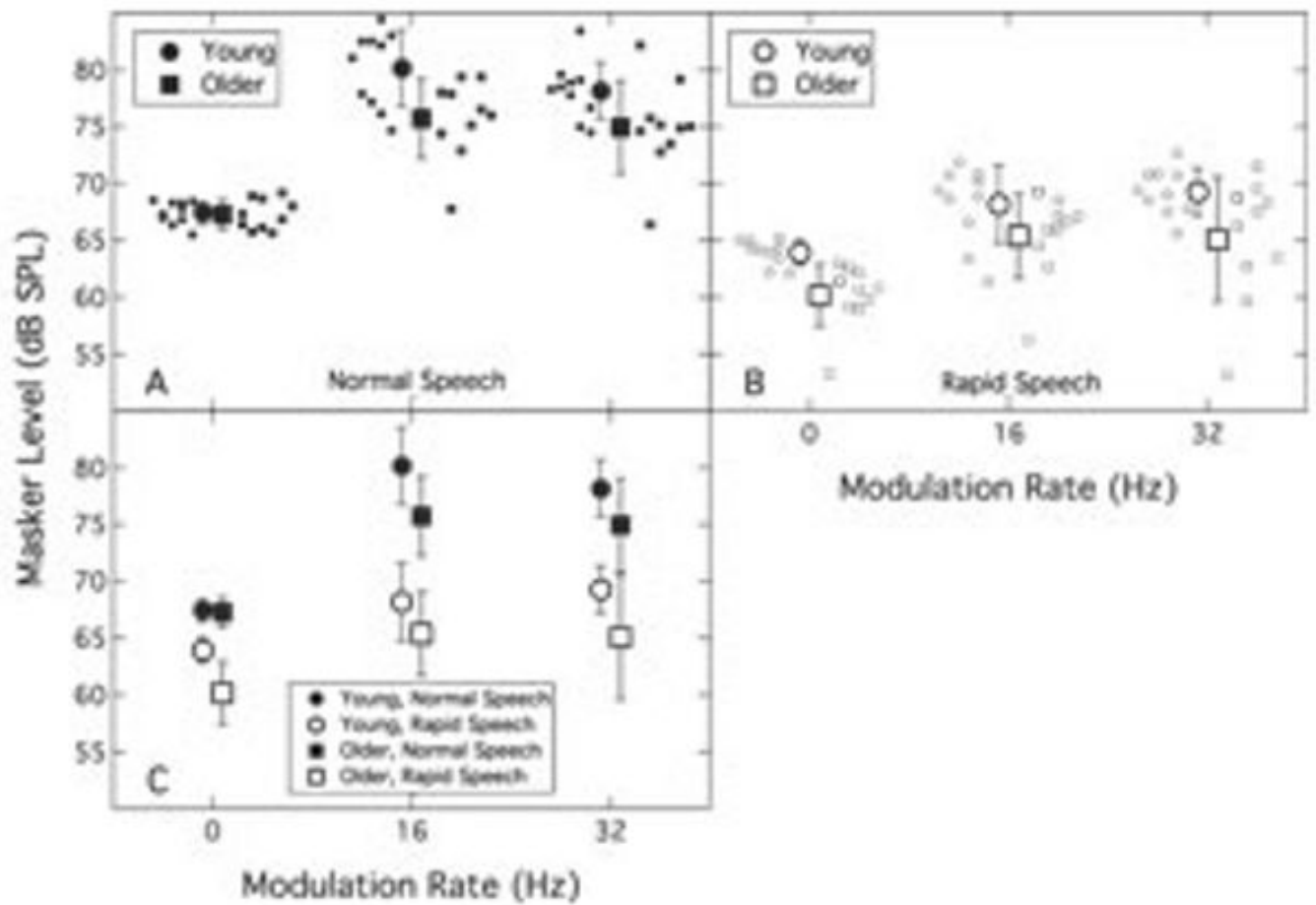
Fig. 1.
Masker level at threshold plotted as a function of masker modulation rate. Panels A and B are individual and mean data for normal and rapid speech, respectively, for both younger (circles) and older (squares) listeners. The mean data and associated standard deviations are re-plotted for all conditions in panel C for both age groups.
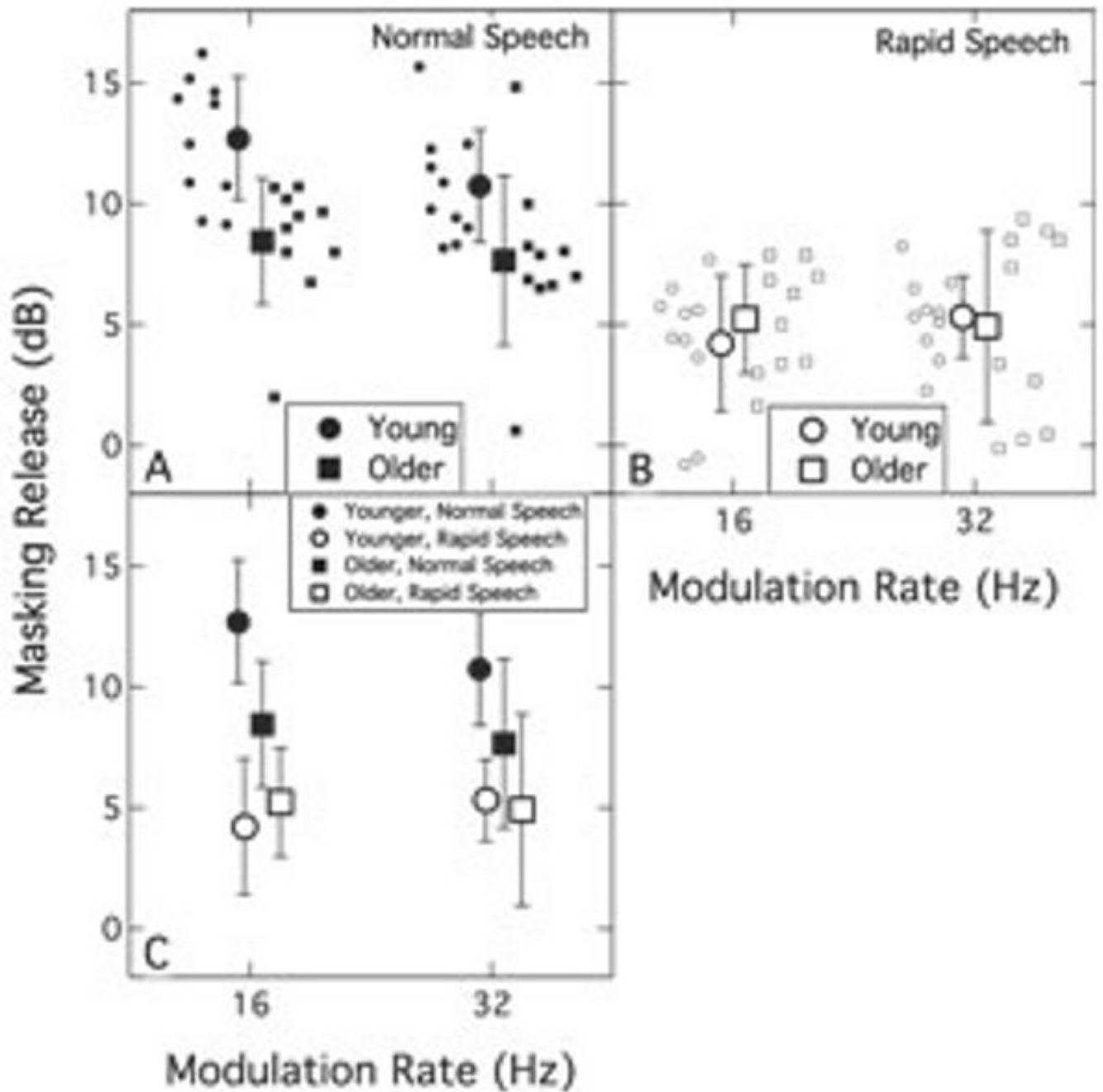
Fig. 2.
Modulation masking release plotted as a function of modulation rate. Panels A and B are individual and mean data for normal and rapid speech, respectively, for both younger (circles) and older (squares) listeners. The mean data and associated standard deviations are re-plotted for all conditions in panel C for both age groups.
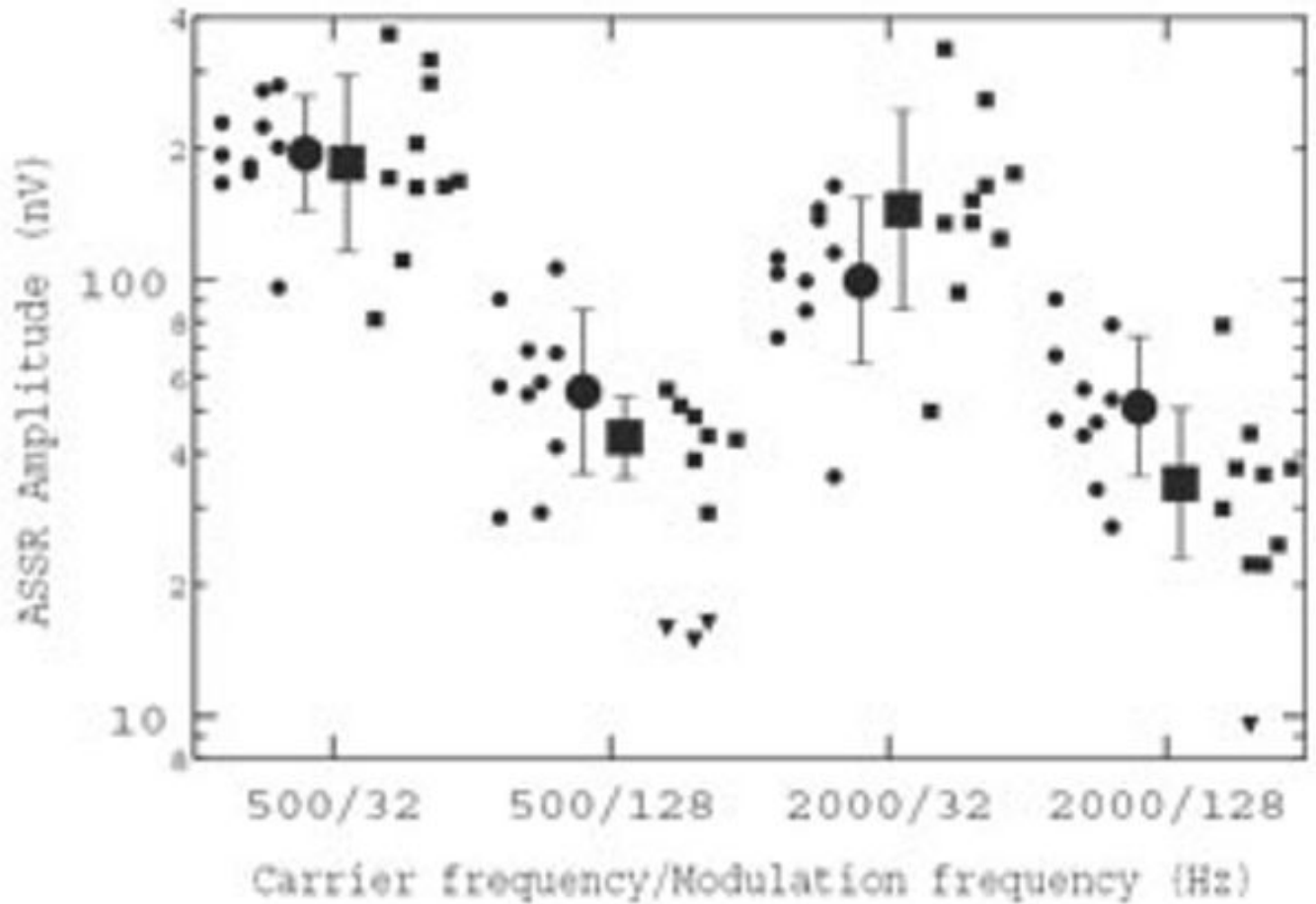
Fig. 3.
ASSR amplitude plotted for each combination of ƒc and ƒm. Mean data, with associated standard deviations, are shown as large symbols, with individual data offset to either side for younger (circles) and older (squares) listeners. For cases (n = 4) where the EEG component at ƒm = 128 Hz was not significantly different from the surrounding EEG noise, the amplitudes are identified as inverted triangles.

Table 1

Mean audiometric thresholds (dB HL) in the test ear for the two age groups. Standard deviations are in parentheses.

	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
Younger	2.50 (6.35)	-0.50 (5.99)	2.50 (6.35)	3.00 (5.87)	1.00 (6.15)	10.00 (7.07)
Older	12.50 (5.89)	10.00 (5.77)	5.50 (6.85)	10.50 (5.50)	14.50 (7.98)	35.71 (10.58)