



Published in final edited form as:

*Ear Hear.* 2019 ; 40(4): 1009–1015. doi:10.1097/AUD.0000000000000683.

## Masking release for speech in modulated maskers: Electrophysiological and behavioral measures

A. Michelle Tanner<sup>1</sup>, Emily R. Spitzer<sup>1,2</sup>, JP Hyzy<sup>1</sup>, and John H. Grose<sup>1</sup>

<sup>1</sup>Department of Otolaryngology - Head & Neck Surgery, University of North Carolina at Chapel Hill

<sup>2</sup>Div. Speech & Hearing Sciences, Department of Allied Health Sciences, University of North Carolina at Chapel Hill

### Abstract

**Objectives**—The purpose of this study was to obtain an electrophysiological analog of masking release using speech-evoked cortical potentials in steady and modulated maskers, and to relate this masking release to behavioral measures for the same stimuli. The hypothesis was that the evoked potentials can be tracked to a lower stimulus level in a modulated masker than in a steady masker, and that the magnitude of this electrophysiological masking release is of the same order as that of the behavioral masking release for the same stimuli.

**Design**—Cortical potentials evoked by an 80-ms /ba/ stimulus were measured in two steady maskers (30- and 65-dB SPL), and in a masker that modulated between these two levels at a rate of 25 Hz. In each masker, a level series was undertaken to determine electrophysiological threshold. Behavioral detection thresholds were determined in the same maskers using an adaptive tracking procedure. Masking release was defined as the difference between signal thresholds measured in the steady 65-dB SPL masker and the modulated masker. A total of 23 normal-hearing adults participated.

**Results**—Electrophysiological thresholds were uniformly elevated relative to behavioral thresholds by about 6.5 dB. However, the magnitude of masking release was about 13.5 dB for both measurement domains.

**Conclusions**—Electrophysiological measures of masking release using speech-evoked cortical auditory evoked potentials correspond closely to behavioral estimates for the same stimuli. This suggests that objective measures based on electrophysiological techniques can be used to reliably gauge aspects of temporal processing ability.

### Introduction

Detection threshold for a signal in a modulated masker is usually lower than it is in a steady masker. This observation has been leveraged into a gauge of temporal resolution known as the masking period pattern [MPP] wherein threshold for a brief tone is measured as a

---

Corresponding author: John H. Grose, Ph.D., Dept OHNS, CB# 7070, University of North Carolina at Chapel Hill, 170 Manning Dr, Chapel Hill, NC 27599-7070, john\_grose@med.unc.edu, Phone: (919) 966-9777.

Conflicts of Interest and Sources of Funding: This research was funded by NIH NIDCD 5-R01-DC001507 (JHG)

function of its temporal position within the modulation pattern of the masker (Zwicker 1976). The more accurately the tone's threshold-by-temporal-position curve (i.e., the MPP) parallels the masker's modulation pattern, the more acute the temporal resolution. Greater susceptibility to temporal masking can lead to a shallower MPP (Zwicker and Schorn 1982). Deficits in temporal resolution, as measured with the MPP, have been observed in older listeners (Grose et al. 2016), in listeners with cochlear hearing loss (Zwicker and Schorn 1982), and in school-age children (Buss et al. 2013). Because the detection of brief signals can be perceptually challenging for children, a coarser measure of temporal processing known as the modified MPP has also been used (Grose et al. 1993). Here, a long-duration signal is used that extends over multiple periods of the masker's modulation. Signal threshold is measured in the steady and modulated maskers, and, in the latter case, without respect to the timing of the signal relative to the modulator phase. The difference in threshold between the two masker conditions is taken as a proxy measure of temporal resolution.

The benefit of a modulated masker for signal detection has also been extensively examined in the context of speech perception (e.g., Bernstein et al. 2012; Desloge et al. 2010; Dirks and Bower 1970; Francart et al. 2011; Fullgrabe et al. 2006; Gnansia et al. 2008; Miller and Licklider 1950; Oxenham and Simonson 2009; Stuart et al. 2006). The speech reception threshold (SRT) is typically lower in a modulated masker than in a steady masker, and this masking release is again thought to depend upon temporal processing ability. However, the advantage to speech perception of listening in a modulated masker is likely multi-faceted, not least because the speech signal being tested can range from simple phonemes to complete sentences. As with signal detection in the MPP, the availability of speech fragments within the masker minima is dependent on temporal masking effects, but the extracted fragments must also be successfully spliced together to reconstruct an intelligible speech signal. A number of studies have examined the factors that affect speech 'glimpsing' (Cooke 2006), but the importance of a temporal processing component is undisputed. A related approach that has also demonstrated the importance of temporal processing to speech perception in modulated noise is the measurement of phoneme recognition in masker minima. Here, an entire consonant-vowel-consonant (CVC) word is placed within a masker minimum and the relative recognition of initial and final consonants assessed as a gauge of the temporal asymmetry of forward and backward masking (Porter et al. 2018). Note that to accommodate an entire CVC word within a masker minimum mandates a slow modulation rate. In summary, masking release for speech in modulated maskers is affected by the fidelity of temporal processing. Diminished masking release has been observed in older listeners (Dubno et al. 2003; Goossens et al. 2017; Grose et al. 2009; Stuart and Phillips 1996), in listeners with cochlear hearing loss (George et al. 2006; Goossens et al. 2017; Jin and Nelson 2006), and in children (Buss et al. 2016). This reduction has been ascribed, at least in part, to deficient (or, in children, immature) temporal processing.

There has been an interest in developing electrophysiological tests of temporal processing that parallel the behavioral tests. Such electrophysiological tests are not only informative as to underlying mechanism(s), but they also have the potential to assess temporal processing abilities in participants who are unable to provide reliable behavioral responses. The objective measures might therefore have clinical relevance. For example, Androulidakis and

Jones (2006) measured cortical auditory evoked potentials (CAEPs) using a paradigm that was analogous to a modified MPP task. They measured the P1-N1-P2 response evoked by a diotic fixed-level, 200-ms tone that was presented in either a steady noise or in a modulated noise. The signal-to-noise ratio (SNR) was set such that the tonal signal was just masked by the steady noise. They observed a robust CAEP in the modulated noise but no response in the steady noise, indicating an ‘unmasking’ of the tone in the modulated noise.

Developing an electrophysiological test of temporal processing that parallels behavioral measures is of particular interest in the context of masking release for speech. Using an envelope following response (EFR) paradigm, Schoof and Rosen (2016) measured the EFR evoked by a diotic 100-ms synthetic vowel /a/ that was masked, in one condition, by a 10-Hz amplitude modulated noise at a fixed SNR. They compared the response evoked by the stimulus segment falling within the masker peak with that of the stimulus segment falling within the masker dip. Differences between these two response regions were taken as a measure of neural masking release. Neural masking release was measured in both younger and older subjects with relatively normal audiograms, and age-related differences were found in the overall robustness of the responses. However, speech perception tests in these same listeners undertaken in steady and modulated maskers showed no age-related differences in SRTs and, therefore, no age-related differences in perceptual masking release. Thus, neural masking release did not parallel perceptual masking release in this study.

A different approach to electrophysiological measures of release from masking using speech stimuli has employed speech-evoked CAEPs measured in modulated and unmodulated noise. Billings et al. (2011) compared P1-N1-P2 responses evoked by monaural presentation of the syllable /ba/ in a steady speech-shaped noise and in the same noise interrupted randomly with silent gaps lasting 5 – 95 ms. A fixed signal level of 65 dB SPL and a fixed SNR of –3 dB was employed. They found no significant differences in response morphology between these two conditions, but pointed out that an electrophysiological masking release might not be observed at this SNR unless the masker interruptions were restricted to gaps longer than about 30 ms. A later study from this laboratory compared CAEPs evoked by the monaural /ba/ stimulus in both a steady speech-shaped noise and in a broadband noise that had been modulated by the envelope of a single talker (Maamor and Billings 2017). Signal level was again fixed at 65 dB SPL, but now three SNRs were tested: –3, 3, and 9 dB. Although response differences were observed between these two maskers, the direction and magnitude of these differences depended on the SNR, possibly due to a floor effect in CAEP morphology measured in the modulated masker. This interaction between masker type and SNR makes a simple interpretation in terms of masking release challenging. More recently, Faucette and Stuart (2017) measured /da/-evoked CAEPs in noise that was either steady or randomly interrupted in the same manner as that noted above in Billings et al. (2011). In one set of monaural conditions, the masker level was fixed at 65 dB SPL and the target speech level was adaptively varied to determine an electrophysiological threshold. The threshold difference between the steady and interrupted masker yielded a masking release of about 17 dB. Thus, several studies have examined speech-evoked CAEPs in modulated and unmodulated noise, but have arrived at disparate conclusions in terms of release from masking. This is likely due, in part, to differences in stimulus SNRs. The challenge of reconciling these disparate conclusions is compounded by the fact that the

electrophysiological measures were not accompanied by parallel behavioral measures. Such parallels are paramount in determining whether electrophysiological measures are predictive of behavioral performance. The likelihood of this is supported by Billings et al. (2015) who showed that, in the absence of hearing loss, monaural /ba/-evoked CAEPs in continuous speech-shaped noise correlated well with measures of speech intelligibility in that noise.

In summary, behavioral studies using both tonal and speech signals have shown masking release associated with a modulated masker. Efforts to demonstrate an electrophysiological parallel to this have been mixed. Tone-evoked CAEPs results have been reported that could be interpreted in terms of an electrophysiological analog of masking release, although these data did not supply an actual measure of masking release magnitude and were not accompanied by complementary behavioral data (Androulidakis and Jones 2006). Speech-evoked CAEPs obtained in steady noise show good correspondence to behavioral measures (Billings et al. 2015), but robust demonstrations of electrophysiological masking release using these speech stimuli in steady and modulated noise have been inconsistent across studies, in part because of the use of limited signal and SNR levels (Billings et al. 2011; Faucette and Stuart 2017; Maamor and Billings 2017). A neural masking release has been demonstrated using the EFR, but observed age effects were not accompanied by corresponding age-related effects in behavioral measures of masking release for speech (Schoof and Rosen 2016). The purpose of the present study, therefore, was to demonstrate an electrophysiological analog of masking release using speech-evoked CAEPs and to relate this to behavioral measures of speech detection for the same stimuli. The hypothesis was that CAEPs can be tracked to a lower stimulus level in a modulated masker than in a steady masker, and that the magnitude of this electrophysiological masking release is of the same order as that of the behavioral masking release for the same stimuli.

## Materials & Methods

### Participants

Participants in the study were normal-hearing adults with audiometric thresholds  $\leq 20$  dB HL across the octave frequencies 250 – 8000 Hz. The median age was 24 years. Twenty-two participants (14 female) undertook conditions in the behavioral task and 18 participants (10 female) undertook conditions in the electrophysiological task; 17 participants undertook conditions in both tasks. Table 1 shows the distribution of participants by task and condition. All participants provided informed consent, and the study was approved by the Institutional Review Board of the University of North Carolina at Chapel Hill. Participants were paid for their participation.

### Stimuli

The signal was the consonant-vowel token /ba/ that originated from the corpus of Stephens and Holt (2011). The digital waveform was up-sampled from the original rate of 11,025 Hz to a rate of 24,414 Hz to be compatible with the RZ6 digital signal processing platform employed in the experiment (Tucker-Davis Technologies [TDT], Alachua, FL). The waveform was also truncated to a duration of 80 ms, beginning and ending at zero crossings. The /ba/ stimulus was calibrated with reference to the dB SPL of a continuous 1-kHz tone

that had the same peak-to-peak amplitude as the /ba/ waveform. The masker was a speech-shaped noise having the same long-term average spectrum of multi-lingual speech defined by Byrne et al. (1994). In two conditions, this noise was presented continuously at a fixed level of either 30 dB SPL or 65 dB SPL (*Steady Low*, *Steady High*). In two other conditions, shown schematically in Figure 1, the noise was quasi square-wave modulated between these two levels, with the level transitions being shaped by 5-ms ramps. In one modulated condition, the rate of modulation was 6.25 Hz (*Mod Slow*). Here, the 160-ms period ensured that the 80-ms /ba/ stimulus was exactly accommodated within the dip of the masker. In the second modulated condition, the modulation rate was 25 Hz (*Mod Fast*). Here, the 40-ms period meant that two sequential masker dips contained segments of the /ba/ stimulus. The stimuli were presented diotically through Etymotic ER2 insert phones (Elk Grove Village, IL). For the electrophysiological section of the study, the insert phones were electromagnetically shielded.

## Procedure

**Behavioral task**—Behavioral thresholds for detecting the /ba/ signal were measured using a 3-alternative, forced-choice procedure that incorporated a 2-down, 1-up stepping rule to converge on the 71% correct point on the psychometric function. In each trial, 400-ms listening intervals – marked by lights on a response box – were separated by 400-ms pauses. Participants entered their response by means of a button press, and correct feedback was given after each trial via the lights. The initial step size for the adaptive procedure was 4 dB, and after the second reversal in level direction this step size was halved to its final step size of 2 dB. A threshold-estimation track was terminated after eight reversals, and the mean signal level at the final six reversal points was taken as the estimate of threshold. At least three threshold estimates were obtained per condition, with a fourth estimate obtained if the range of the first three exceeded 3 dB. All estimates were averaged to yield the final threshold for that condition. Initially, behavioral thresholds were measured for three conditions: *Steady Low*, *Steady High*, and *Mod Slow*. Thirteen subjects participated in these initial conditions. Later, the *Mod Fast* condition was added, and 11 subjects participated in this condition, two of whom had participated in the initial conditions.

**Electrophysiological task**—Participants were asked to sit quietly in a comfortable chair and remain alert by watching a silent, subtitled movie or by reading. Silver-silver chloride electrodes were placed in a two-channel montage with the inverting electrodes positioned on the left and right earlobes (A1 and A2, respectively). The non-inverting electrode for both channels was initially positioned at the high forehead hairline (~Fpz) for the conditions *Steady Low*, *Steady High*, and *Mod Slow*, and was then moved to vertex (Cz) when the *Mod Fast* condition was added to replace *Mod Slow* which exhibited masker contamination (see below). Ground was placed at low forehead. Impedance values were maintained  $< 5$  k $\Omega$ , with impedances matched within 3 k $\Omega$  across electrodes. Electroencephalographic (EEG) activity was continuously recorded with a Neuroscan SynAmpRT system (Compumed, Los Angeles, CA) operating at a sampling rate of 1000 Hz. Artifact rejection was set at  $\pm 70$   $\mu$ V, and the recordings were filtered from 1–100 Hz. The Neuroscan recording system was synchronized to the TDT stimulation system by means of a time-event marker ('trigger') coincident with the initiation of each /ba/ stimulus. Relative to this onset time marker, EEG responses were

segmented into epochs extending from –100 ms to +300 ms. A recording was terminated after 200 artifact-free epochs were collected for each condition. Offline, the recordings were baseline corrected and a linear transformation applied to derive a single vertical channel referenced to both earlobes. The 200 epochs were averaged and re-filtered from 1–30 Hz to smooth the response.

For each masker condition (*Steady Low*, *Steady High*, *Mod Slow*, and *Mod Fast*) a level series of recordings evoked by the /ba/ stimulus was undertaken. The range of levels used for each masker condition was dictated by the group mean behavioral threshold for that condition, and typically ranged from approximately 0 – 40 dB SL re that threshold in 5- or 10-dB steps. For each averaged recording, a response was determined to be present based on some combination of the following four criteria: (1) a root-mean-square (RMS) voltage in the post-stimulus region of interest (ROI) that was  $\geq 50\%$  re the RMS in a defined pre-stimulus window; relative to the onset time marker, the ROI extended from +50 ms to +300 ms, and the pre-stimulus baseline extended from –300 ms to –50 ms; (2) a P1-N1 amplitude  $\geq 1.5 \mu\text{V}$ ; (3) a N1-P2 amplitude  $\geq 1.5 \mu\text{V}$ ; and (4) visual confirmation of a present response by two experts in cortical potential waveform analysis. For 73 % of responses that were deemed present, all four criteria were met. For 22 % of responses deemed present, three of the four criteria were met. For the remaining 5 % of the responses deemed present, only two criteria were met and in these cases agreement between the two expert judges was given the greater weight. Of the 18 participants who provided data for the electrophysiological task, 10 participants (5 female) received the initial *Mod Slow* condition. Following the introduction of the *Mod Fast* condition, data from a further 10 participants (6 female) were collected; two participants provided data for both the *Mod Slow* and *Mod Fast* conditions.

## Results & Discussion

### Behavioral task

The group mean thresholds for the /ba/ stimulus as a function of masker type are shown in Figure 2. Signal thresholds in the *Steady Low* and *Steady High* maskers correspond closely to the actual masker levels for those two conditions (30 dB SPL and 65 dB SPL). Threshold in the *Mod Slow* condition, where the entire /ba/ stimulus was contained within a masker dip, is only about 4 dB higher than the masker floor. This suggests that detection of the /ba/ stimulus during the relatively long dip was resilient to temporal masking by the surrounding masker peaks. In contrast, the 50-dB SPL signal threshold in the *Mod Fast* condition is markedly elevated above the masker floor, indicating that temporal masking effects were much more robust in this condition. Irrespective of the differences in mean threshold across the four conditions, there was very little variability in threshold across participants within a condition, as indicated by the small standard deviations. The difference in mean thresholds between the *Steady High* and *Mod Fast* conditions was 13.3 dB. An independent-samples t-test showed this difference to be significant ( $t[22] = 25.2$ ;  $p < 0.001$ ).

### Electrophysiological task

The response waveforms for the /ba/ level series in the *Steady Low*, *Steady High*, and *Mod Fast* maskers are shown in the left, middle, and right panels, respectively, of Figure 3. Note

the different intensity ranges for each panel. The individual traces are shown as light grey lines and the group mean responses as heavy dark lines. These group mean response waveforms are replotted in panels A – C of Figure 4, in collapsed format. This format draws attention to the observation that as the SNR decreases, the overall amplitude of the P1-N1-P2 response declines and the peak latencies increase – most notably that of N1. The similarity of the patterns across the two steady maskers, although collected at very different overall levels, highlights the dependency of response amplitude on SNR rather than on absolute signal level, in confirmation of the report by Billings et al. (2009). This similarity was confirmed with a linear mixed model analysis comparing the P1-N1 amplitudes across the two steady maskers, selecting levels 40, 45 and 50 dB SPL for the *Steady Low* masker, and 70, 75, and 80 dB SPL for the *Steady High* masker. The analysis showed a significant effect of stimulus intensity ( $F[2,27.344] = 7.203$ ;  $p = 0.003$ ) but no effect of overall masker level ( $F[1,36.972] = 1.875$ ;  $p = 0.179$ ). A similar analysis on N1 latency at these levels showed a significant effect of stimulus intensity ( $F[2,27.817] = 6.609$ ,  $p = 0.004$ ) as well as a significant effect of overall masker level ( $F[1,41.616] = 5.844$ ;  $p = 0.02$ ), with latencies being longer at the higher masker level.

Using the same collapsed format, the group mean response waveforms in the *Mod Slow* masker are shown in panel D of Figure 4. These traces show a repeatable aberration in the pre-stimulus period, as indicated by an arrow. This artifact is due to the averaging system being time-locked not only to the onset of the /ba/ stimulus but also to the surrounding level transitions in the modulated masker. Although the 6.25-Hz modulations of the *Mod Slow* masker occur more frequently than the stimulus, the rate is apparently slow enough that masker-generated responses remain sufficiently resilient to adaptation that they contaminate the recording. Masker-generated responses are also evident in the time-locked data of Androulidakis and Jones (2006) who used a 17.5-Hz modulation rate. It was this masker contamination of the responses that prompted the change in modulation rate from 6.25 Hz to 25 Hz. At the more rapid modulation rate, neural adaptation to the masker modulations is sufficient to mitigate the masker-related artifact in the response, as shown by the absence of a time-locked response at negative latencies in the *Mod Fast* recordings (panel C).

### Comparison of Behavioral and Electrophysiological tasks

Of key interest is whether an electrophysiological analog of masking release is present. To quantify this, the electrophysiological threshold for each participant was determined for the *Steady High* and *Mod Fast* conditions. For each level series, the electrophysiological threshold was taken to be the lowest level at which a response could be detected. The group mean electrophysiological thresholds for the two conditions are shown as dark bars in Figure 5. The difference in mean thresholds between the *Steady High* and *Mod Fast* conditions is 13.5 dB. For comparison, the mean behavioral thresholds for these two conditions are shown as light bars. An analysis of variance (ANOVA) on these data indicated a significant effect of masker condition (*Steady High*, *Mod Fast*) [ $F(1,40) = 327.19$ ;  $p < 0.001$ ] and a significant effect of task (electrophysiological, behavioral) [ $F(1,40) = 76.58$ ;  $p < 0.001$ ], but no interaction between these two factors [ $F(1,40) = 0.012$ ;  $p = 0.91$ ]. The absence of an interaction signifies that the 13.5-dB masking release seen in the electrophysiological data does not differ from the 13.3-dB masking release seen in the behavioral data.

The data in Figure 5 suggest that there is a good correspondence between the electrophysiological and behavioral thresholds. To demonstrate this further, the individual electrophysiological and behavioral thresholds in the *Steady Low*, *Mod Fast*, and *Steady High* conditions are plotted as a scattergram in Figure 6. Although thresholds did not vary greatly across participants within a condition, particularly for the behavioral thresholds (cf Figure 2), the scattergram confirms that there was generally a close association between the electrophysiological and behavioral thresholds across conditions. The electrophysiological threshold exceeded the behavioral threshold by, on average, 6.4 dB for *Steady Low*, 6.4 dB for *Mod Fast*, and 6.6 dB for *Steady High*. This offset, or ‘correction factor’, corresponds precisely with the 6.5-dB correction factor derived by Lightfoot and Kennedy (2006) for CAEP estimates of behavioral threshold *in quiet*, and is with the range of 1 – 12 dB found by other studies (e.g., Tomlin et al. 2006; Tsu et al. 2002; Yeung and Wong 2007). This suggests that speech-evoked CAEP thresholds can be reliable predictors of speech detection thresholds in both steady and fluctuating maskers. This is promising for their use in the clinical assessment of patients who are unable to provide reliable behavioral responses, such as infants. Because speech *intelligibility* was not assessed in this study, no comment can be made on the ability of speech-evoked CAEPs to predict SRTs or other measures of speech recognition. However, other studies have shown a strong predictive association between similar CAEP measures and speech perception (e.g., Billings et al. 2015).

## Summary & Conclusions

The purpose of this study was to compare electrophysiological and behavioral measures of release from masking for speech sounds. Thresholds for speech-evoked CAEPs were measured in steady and modulated noise in order to derive a neural masking release, and behavioral thresholds for the same stimuli were also determined. The results showed that, in conditions where the /ba/-evoked CAEP response was not contaminated by masker-evoked response, the CAEP threshold corresponded well to that of the behavioral threshold for the same condition. The electrophysiological threshold was uniformly about 6.5 dB higher than the behavioral threshold. For both electrophysiological and behavioral measures, the average threshold in the modulated masker about 13.5 dB lower than that in the steady masker, yielding a derived masking release that was similar in both domains. This similarity in performance across electrophysiological and behavioral domains underscores the benefit of objective measures that parallel perception and that, therefore, may have applicability in testing populations unable to provide reliable behavioral results. The findings of the present study are limited predominantly to young, normal-hearing adults. Because behavioral studies have shown deficits in MPPs at both ends of the age span (Buss et al. 2013; Grose et al. 2016), and electrophysiological studies have shown age-related differences in speech-evoked CAEPs at both ends of the age span (Maamor and Billings 2017; O’Brien et al. 2015; Wunderlich and Cone-Wesson 2006), the next step in this investigation is to assess the comparative effects of age on behavioral and electrophysiological measures of release from masking for speech sounds.



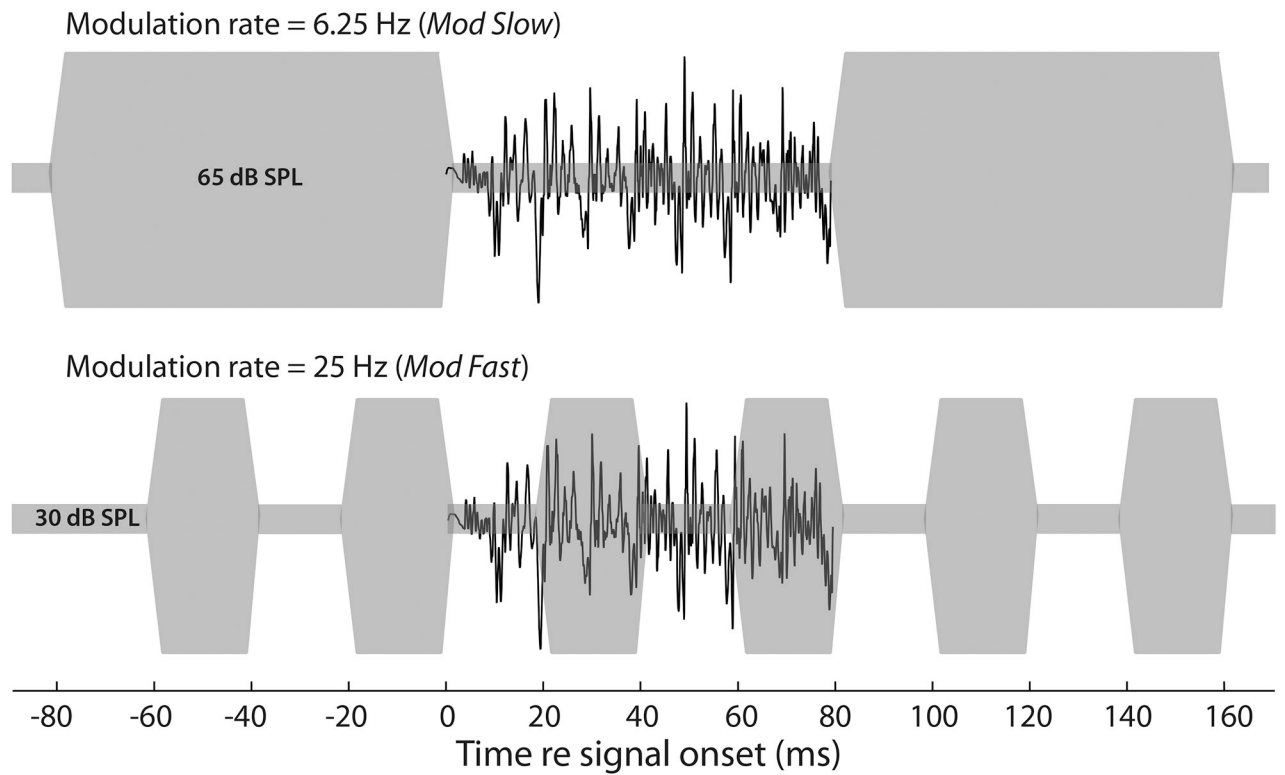
## Acknowledgements

All authors contributed to multiple aspects of the study. Broadly, ERS and JHG designed and implemented the experiments, while AMT and JH performed the experiments and contributed to data analysis. All authors assisted in manuscript preparation.

## References

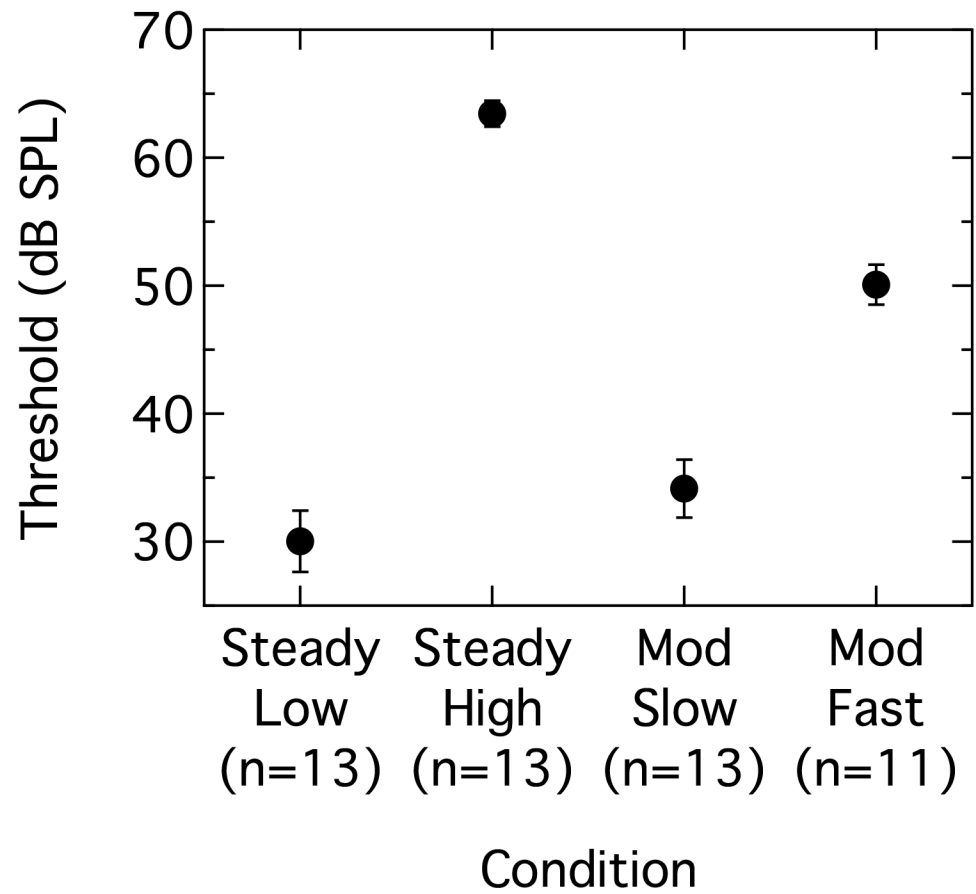
- Androulidakis AG, Jones SJ (2006). Detection of signals in modulated and unmodulated noise observed using auditory evoked potentials. *Clin Neurophysiol*, 117, 1783–1793. [PubMed: 16793334]
- Bernstein JG, Summers V, Iyer N, et al. (2012). Set-size procedures for controlling variations in speech-reception performance with a fluctuating masker. *J Acoust Soc Am*, 132, 2676–2689. [PubMed: 23039460]
- Billings CJ, Bennett KO, Molis MR, et al. (2011). Cortical encoding of signals in noise: effects of stimulus type and recording paradigm. *Ear Hear*, 32, 53–60. [PubMed: 20890206]
- Billings CJ, Penman TM, McMillan GP, et al. (2015). Electrophysiology and perception of speech in noise in older listeners: Effects of hearing impairment and age. *Ear Hear*, 36, 710–722. [PubMed: 26502191]
- Billings CJ, Tremblay KL, Stecker GC, et al. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hear Res*, 254, 15–24. [PubMed: 19364526]
- Buss E, He S, Grose JH, et al. (2013). The monaural temporal window based on masking period pattern data in school-aged children and adults. *J Acoust Soc Am*, 133, 1586–1597. [PubMed: 23464028]
- Buss E, Leibold LJ, Hall JW 3rd. (2016). Effect of response context and masker type on word recognition in school-age children and adults. *J Acoust Soc Am*, 140, 968. [PubMed: 27586729]
- Byrne D, Dillon H, Tran K, et al. (1994). An international comparison of long-term average speech spectra. *J Acoust Soc Am*, 96, 2108–2120.
- Cooke M (2006). A glimpsing model of speech perception in noise. *J Acoust Soc Am*, 119, 1562–1573. [PubMed: 16583901]
- Desloge JG, Reed CM, Braida LD, et al. (2010). Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise. *J Acoust Soc Am*, 128, 342–359. [PubMed: 20649229]
- Dirks D, Bower D (1970). Effects of forward and backward masking on speech intelligibility *J Acoust Soc Am*, 47, 1003–1008. [PubMed: 5443146]
- Dubno JR, Horwitz AR, Ahlstrom JB (2003). Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing. *J Acoust Soc Am*, 113, 2084–2094. [PubMed: 12703719]
- Faucette SP, Stuart A (2017). Evidence of a speech evoked electrophysiological release from masking in noise. *J Acoust Soc Am*, 142, EL218. [PubMed: 28863590]
- Francart T, van Wieringen A, Wouters J (2011). Comparison of fluctuating maskers for speech recognition tests. *Int J Audiol*, 50, 2–13. [PubMed: 21091261]
- Fullgrave C, Berthommier F, Lorenzi C (2006). Masking release for consonant features in temporally fluctuating background noise. *Hear Res*, 211, 74–84. [PubMed: 16289579]
- George EL, Festen JM, Houtgast T (2006). Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, 120, 2295–2311. [PubMed: 17069325]
- Gnansia D, Jourdes V, Lorenzi C (2008). Effect of masker modulation depth on speech masking release. *Hear Res*, 239, 60–68. [PubMed: 18434049]
- Goossens T, Vercammen C, Wouters J, et al. (2017). Masked speech perception across the adult lifespan: Impact of age and hearing impairment. *Hear Res*, 344, 109–124. [PubMed: 27845259]
- Grose JH, Hall JW, Gibbs C (1993). Temporal analysis in children. *J Speech Hear Res*, 36, 351–356. [PubMed: 8487526]

- Grose JH, Mamo SK, Hall JW 3rd. (2009). Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear Hear*, 30, 568–575. [PubMed: 19633565]
- Grose JH, Menezes DC, Porter HL, et al. (2016). Masking period patterns and forward masking for speech-shaped noise: Age-related effects. *Ear Hear*, 37, 48–54. [PubMed: 26230495]
- Jin SH, Nelson PB (2006). Speech perception in gated noise: the effects of temporal resolution. *J Acoust Soc Am*, 119, 3097–3108. [PubMed: 16708964]
- Lightfoot G, Kennedy V (2006). Cortical electric response audiometry hearing threshold estimation: accuracy, speed, and the effects of stimulus presentation features. *Ear Hear*, 27, 443–456. [PubMed: 16957496]
- Maamor N, Billings CJ (2017). Cortical signal-in-noise coding varies by noise type, signal-to-noise ratio, age, and hearing status. *Neurosci Lett*, 636, 258–264. [PubMed: 27838448]
- Miller GA, Licklider JCR (1950). The intelligibility of interrupted speech. *J Acoust Soc Am*, 22, 167–173.
- O'Brien JL, Nikjeh DA, Lister JJ (2015). Interaction of musicianship and aging: A comparison of cortical auditory evoked potentials. *Behav Neurol*, 2015, 545917. [PubMed: 26504354]
- Oxenham AJ, Simonson AM (2009). Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference. *J Acoust Soc Am*, 125, 457–468. [PubMed: 19173431]
- Porter HL, Spitzer ER, Buss E, et al. (2018). Forward and backward masking of consonants in school-age children and adults. *J Speech Lang Hear Res*, 61, 1807–1814. [PubMed: 29971342]
- Schoof T, Rosen S (2016). The role of age-related declines in subcortical auditory processing in speech perception in noise. *J Assoc Res Otolaryngol*, 17, 441–460. [PubMed: 27216166]
- Stephens JD, Holt LL (2011). A standard set of American-English voiced stop-consonant stimuli from morphed natural speech. *Speech Commun*, 53, 877–888. [PubMed: 21666844]
- Stuart A, Givens GD, Walker LJ, et al. (2006). Auditory temporal resolution in normal-hearing preschool children revealed by word recognition in continuous and interrupted noise. *J Acoust Soc Am*, 119, 1946–1949. [PubMed: 16642807]
- Stuart A, Phillips DP (1996). Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbycusis listeners. *Ear Hear*, 17, 478–489. [PubMed: 8979036]
- Tomlin D, Rance G, Graydon K, et al. (2006). A comparison of 40 Hz auditory steady-state response (ASSR) and cortical auditory evoked potential (CAEP) thresholds in awake adult subjects. *Int J Audiol*, 45, 580–588. [PubMed: 17062499]
- Tsu B, Wong LL, Wong EC (2002). Accuracy of cortical evoked response audiometry in the identification of non-organic hearing loss. *Int J Audiol*, 41, 330–333. [PubMed: 12353605]
- Wunderlich JL, Cone-Wesson BK (2006). Maturation of CAEP in infants and children: a review. *Hear Res*, 212, 212–223. [PubMed: 16480841]
- Yeung KN, Wong LL (2007). Prediction of hearing thresholds: comparison of cortical evoked response audiometry and auditory steady state response audiometry techniques. *Int J Audiol*, 46, 17–25. [PubMed: 17365051]
- Zwicker E (1976). Masking period patterns of harmonic complex tones. *J Acoust Soc Am*, 60, 429–439. [PubMed: 993466]
- Zwicker E, Schorn K (1982). Temporal resolution in hard of hearing patients. *Audiol*, 21, 474–492.

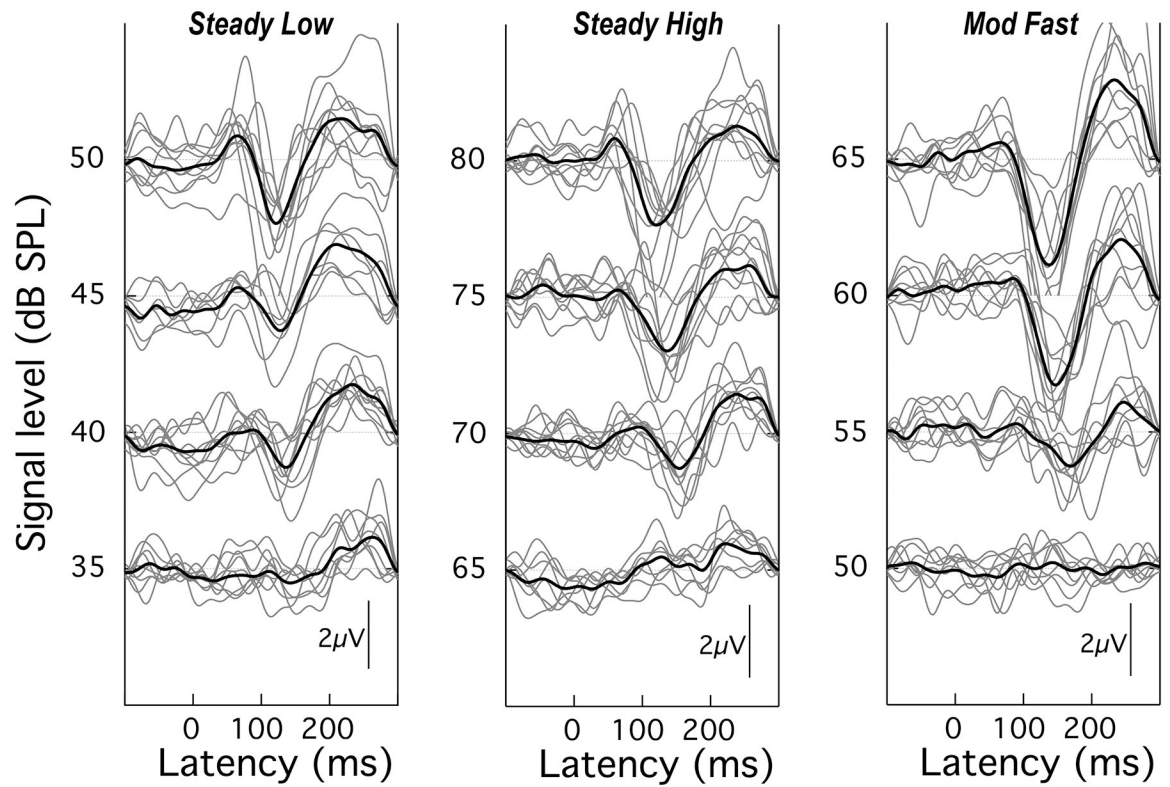


**Figure 1.**

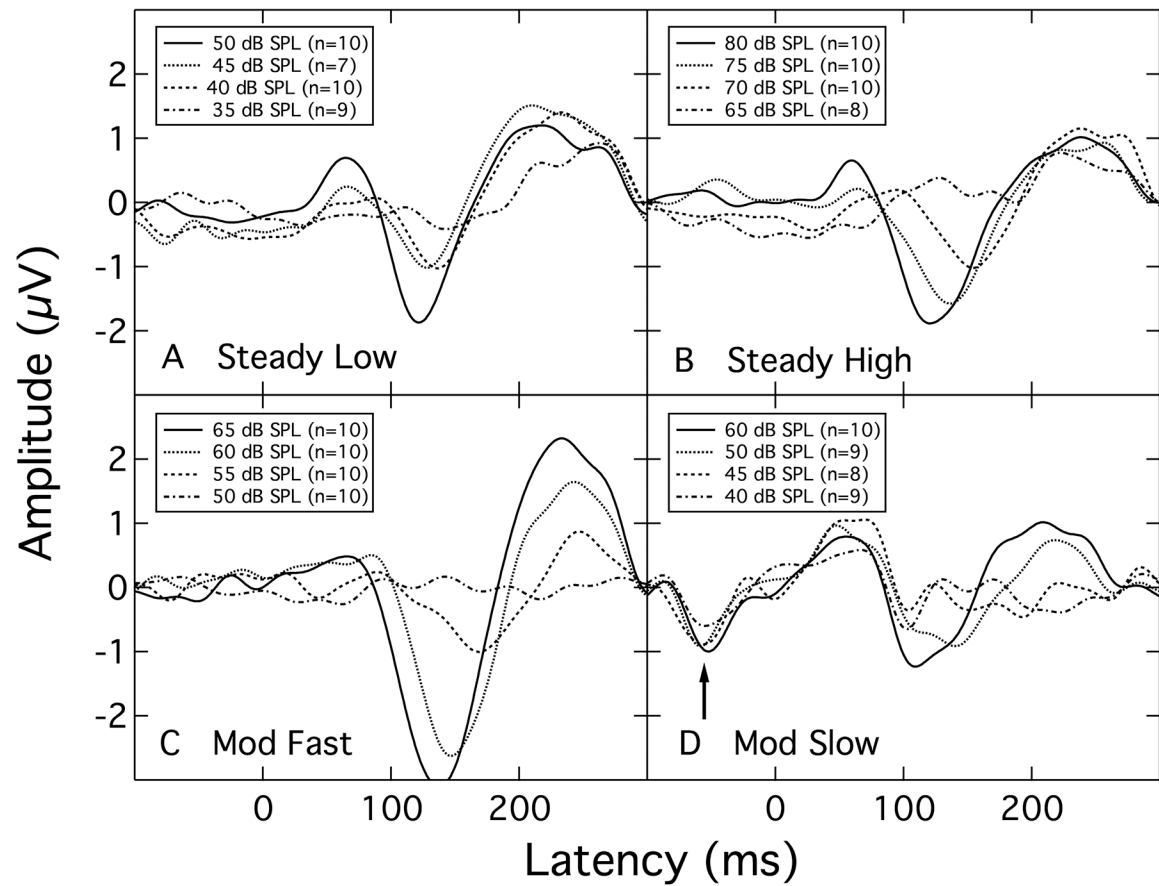
Schematic of the *Mod Slow* (upper trace) and *Mod Fast* (lower trace) conditions. The masker modulated between a level of 65 dB SPL and 30 dB SPL in a quasi square-wave function with 5-ms transition slopes. The /ba/ signal was temporally positioned at the offset of a masker modulation.



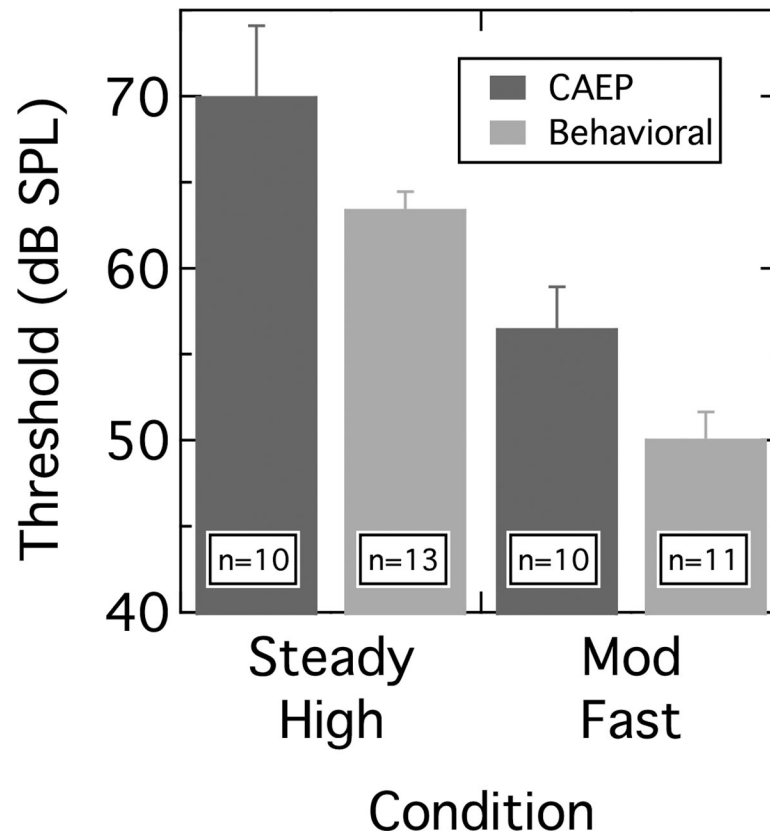
**Figure 2.** Group mean behavioral thresholds for the /ba/ stimulus as a function of masker type. Abscissa notes participant numbers for each condition. Error bars are  $\pm 1$  standard deviation.



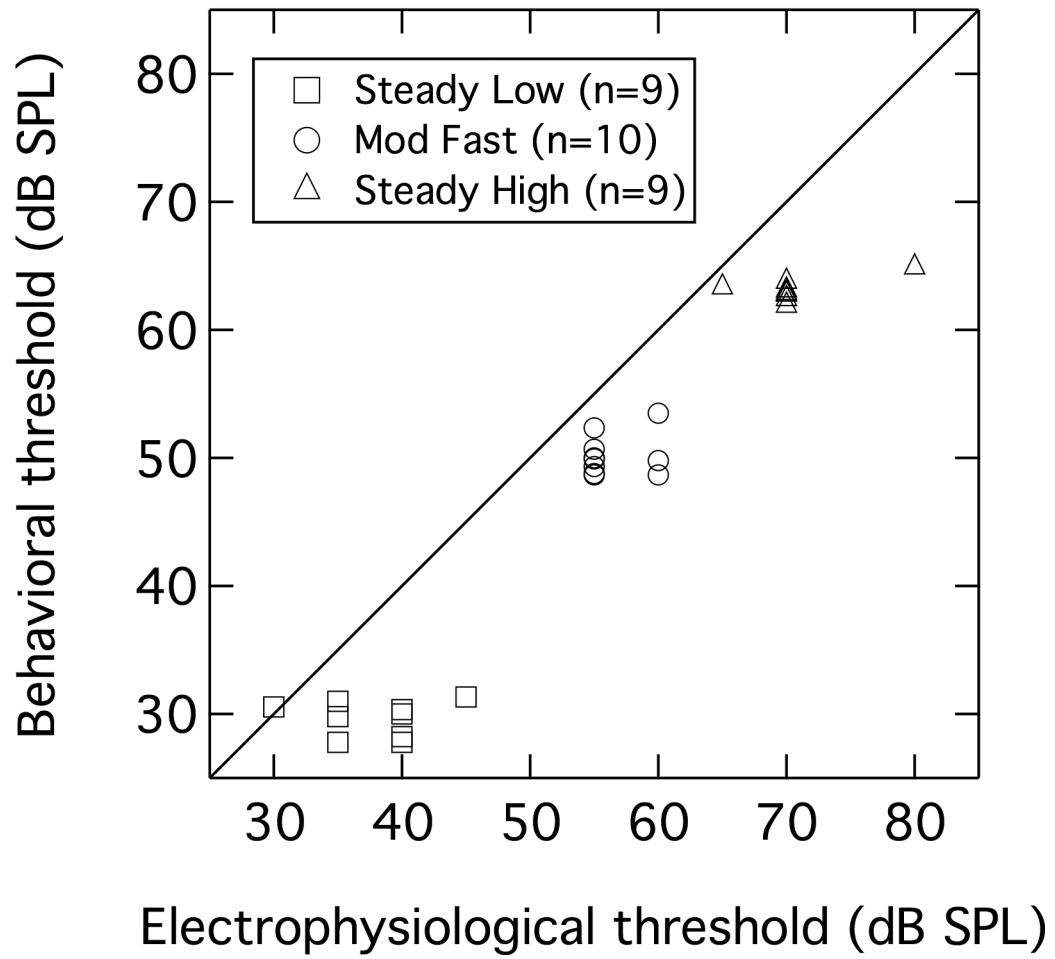
**Figure 3.** Individual (light gray) and group mean (black) CAEPs for the *Steady Low*, *Steady High*, and *Mod Fast* maskers (left to right, respectively). Each panel shows a level series, with a different range of intensities for each panel.



**Figure 4.** Group mean CAEPs for the *Steady Low*, *Steady High*, *Mod Fast*, and *Mod Slow* maskers (Panels A – D, respectively). The arrow in Panel D points to the masker artifact (see text).



**Figure 5.** Mean CAEP thresholds (dark bars) and behavioral thresholds (light bars) in the *Steady High* and *Mod Fast* maskers. Participant numbers for each condition are noted as insets. Error bars are 1 SD.



**Figure 6.** Scattergram of electrophysiological thresholds and behavioral thresholds. The diagonal denotes equivalence.



