Research Article

Effects of Masker Envelope Fluctuations on the Temporal Effect

SKYLER G. JENNINGS,^{1,2} KAYLA SIVAS,¹ AND CAITLIN STONE¹

¹*Department of Communication Sciences and Disorders, The University of Utah, 390 South, 1530 East, BEHS 1201, Salt Lake City, UT 84112, USA*

²*Department of Otolaryngology-Head and Neck Surgery, Medical University of South Carolina, 135 Rutledge Avenue, MSC 550, Charleston, SC 29425-5500, USA*

Received: 29 September 2017; Accepted: 26 June 2018; Online publication: 20 August 2018

ABSTRACT

Under certain conditions, detection thresholds in simultaneous masking improve when the onset of a short sinusoidal probe is delayed from the onset of a long masker. This improvement, known as the temporal effect, is largest for broadband maskers and is smaller or absent for narrowband maskers centered on the probe frequency. This study tests the hypothesis that small or absent temporal effects for narrowband maskers are due to the inherent temporal envelope fluctuations of Gaussian noise. Temporal effects were measured for narrowband noise maskers with fluctuating (“fluctuating maskers”) and flattened (“flattened maskers”) temporal envelopes as a function of masker level (Exp. I) and in the presence of fluctuating and flattened precursors (Exp. II). The temporal effect was absent for fluctuating narrowband maskers and as large as ~7 dB for flattened narrowband maskers. The AC-coupled power of the temporal envelopes of precursors and maskers accounted for 94 % of the variance in probe detection thresholds measured with fluctuating and flattened precursors and maskers. These results suggest that masker temporal envelope fluctuations contribute to the temporal effect and should be considered in future modeling efforts.

Keywords: masking, temporal envelope, temporal effects

INTRODUCTION

In subjects with normal hearing, detection thresholds for a short (e.g., < 20 ms) sinusoidal probe improve as the probe’s onset is delayed several tens of milliseconds from the onset of a longer masker (e.g., > 200 ms). This improvement occurs for broadband (Zwicker 1965a) and off-frequency maskers (McFadden 1989; Schmidt and Zwicker 1991), but not for narrowband maskers centered on the probe frequency (Bacon and Smith 1991; Bacon and Liu 2000). The difference in detection thresholds for probes presented near the temporal center, compared to the onset of the masker, has been referred to as overshoot (Zwicker 1965b), or the temporal effect (Wright 1997). This difference is reported as a positive number and indicates the improvement gained in probe detectability as a result of moving the probe from the onset to the temporal center of the masker. Subjects with permanent cochlear hearing loss (Bacon and Takahashi 1992; Strickland and Krishnan 2005; Jennings et al. 2016) and temporary cochlear hearing loss induced by aspirin ingestion (McFadden and Champlin 1990) or noise exposure (Champlin and McFadden 1989) have smaller temporal effects compared to normal hearing subjects. These smaller temporal effects are due to better-than-normal detection thresholds for probes presented near the masker’s onset. These effects of cochlear hearing loss have led to the hypothesis that the temporal effect depends on the degree of cochlear compression (von Klitzing and Kohlrausch 1994;

Portions of this research were presented at the Thirty-eighth Annual Mid-winter Meeting of the Association for Research in Otolaryngology (2015)

Correspondence to: Skyler G. Jennings · Department of Communication Sciences and Disorders · The University of Utah · 390 South, 1530 East, BEHS 1201, Salt Lake City, UT 84112, USA.
email: skyler.jennings@hsc.utah.edu

Strickland 2001; Bacon and Savel 2004), which is determined by the status of the cochlear amplifier (Davis 1983).

Auditory models have been effective at predicting the temporal effect in subjects with normal (Strickland 2001, 2004) and impaired (Jennings et al. 2011; Jennings et al. 2016) hearing by assuming that (1) better-than-normal thresholds in subjects with hearing impairment for probes presented near the masker's onset are accounted for by less compressive cochlear responses; and (2) the improvement in detection thresholds for probes presented near the temporal center, compared to the onset of maskers in subjects with normal hearing, is accounted for by a decrease in compression over the first several tens of milliseconds of the masker, as a result of a reduction in cochlear amplifier gain (Schmidt and Zwicker 1991; von Klitzing and Kohlrausch 1994). In reference to the latter point, efferent feedback from the medial olivocochlear (MOC) reflex has often been cited as a potential mechanism for the putative reduction in cochlear gain simulated by studies modeling the temporal effect (Strickland 2008; Jennings et al. 2011).

Many of the “gain reduction” models used to predict the temporal effect (Strickland 2001; Bacon and Savel 2004; Strickland 2008; Jennings et al. 2011; Jennings et al. 2016) are modifications of Fletcher's power spectrum model of masking (Fletcher 1940), where detection is based on a constant signal-to-noise ratio (SNR) between the average power of the masker and the average power of the probe+masker at the output of an auditory filter centered on the probe frequency. Modifications required to predict the temporal effect include the simulation of cochlear compression and of a reduction in cochlear gain over the time course of the masker. In addition to predicting the temporal effect in simultaneous masking, cochlear gain reduction models account for similar “temporal effects” in forward masking (Krull and Strickland 2008; Jennings et al. 2009; Roverud and Strickland 2010; Jennings and Strickland 2010, 2012), intensity discrimination (Roverud and Strickland 2015), and amplitude modulation detection (Almishaal et al. 2017). An interpretation of the temporal effect in terms of the status of the cochlear amplifier is significant because it suggests that the normal auditory system, perhaps via efferent feedback, adapts to the local sound environment to improve performance in perceptual tasks, including listening to speech in noisy environments (Guinan 2010; Jennings et al. 2016).

Despite the ecological appeal of interpreting the temporal effect in terms of cochlear amplifier status, such an interpretation cannot fully explain the dependence of the temporal effect on the spectral properties of the masker. For example, it is unclear

why a temporal effect is observed for broadband noise (BBN) maskers (Zwicker 1965b) and tonal maskers centered on the probe frequency (Bacon and Viemeister 1985), but not for narrowband noise (NBN) maskers centered on the probe frequency with bandwidths $\leq 10\%$ of the probe frequency (Bacon and Smith 1991). Opposite of that observed with broadband and tonal maskers, thresholds for short probes are often poorer when presented near the temporal center of NBN maskers, than for probes presented at the masker's onset (Bacon and Smith 1991; Bacon and Liu 2000). This “negative” temporal effect for NBN maskers may be related to differences in the temporal envelopes of NBN maskers, compared to BBN or tonal maskers. The average duration between adjacent temporal envelope dips is relatively longer for NBN compared to that for BBN. Thus, the temporal envelope of NBN fluctuates more slowly over time. In addition to these slower fluctuations, crest factors computed on the noise temporal envelope are larger for NBN than for BBN when temporal envelopes are extracted via rectification and low-pass filtering (e.g., Viemeister 1979). Several models of tone-in-noise detection have shown that a decision variable based on temporal envelope characteristics is consistent with empirical data from human listeners (Richards 1992; Ewert and Dau 2000; Davidson et al. 2006). The question addressed in this study is whether the lack of a temporal effect for NBN maskers is partially due to the inherent temporal envelope fluctuations of these maskers, and if so, to what extent thresholds can be quantitatively accounted for by these fluctuations. Two experiments were conducted to answer these questions. Experiment I involved measuring the temporal effect as a function of probe level for NBN maskers with either fluctuating, or flattened temporal envelopes. Experiment II addressed the influence of temporal envelope fluctuations on masking by measuring probe detection thresholds in the presence of fluctuating and flattened precursors and maskers.

METHODS

Experiments and Subjects

Young normal-hearing adults served as subjects in all experiments. Subjects gave consent and were given a hearing test prior to enrollment. All subjects had audiometric thresholds ≤ 15 dB HL at octave frequencies between 250 and 8000 Hz and normal middle ear status confirmed by otoscopic examination and tympanometry (226-Hz probe tone). Subjects were paid for their participation. Six subjects (S1–S6; ages 25–31 years, 2 male) participated in Exp. 1, and

five subjects (S1, S7–S10; ages 24–31, 1 male) participated in Exp. 2.

Stimuli and Apparatus

The 4000-Hz probe was 6 ms in duration with 3-ms $\cos^2$ rise/fall ramps (no steady-state portion). The probe's onset occurred 2 or 197 ms after the onset of the 400-ms masker, which also had 3-ms $\cos^2$ rise/fall ramps. The bandwidth of the masker was 0.5 equivalent rectangular bandwidths for normal hearing (ERB_N) (228 Hz, Glasberg and Moore 1990), and its spectrum was linearly centered on the probe frequency. Maskers with fluctuating temporal envelopes (fluctuating maskers) were generated by filtering Gaussian white noise, while maskers with flattened temporal envelopes (flattened maskers) were generated by iteratively (10 times) dividing filtered noise by the Hilbert envelope, as described by Kohlrausch et al. (1997). All noises were independently generated (i.e., frozen noise was not used). Schematic representation of the spectral properties of the stimuli is provided in the top panel of Fig. 1. Also, plotted in Fig. 1 is the temporal envelope power at the output of a 1- ERB_N wide gammatone filter centered on the probe frequency for fluctuating and flattened maskers (bottom panel). As shown by waveform insets in Fig. 1, temporal envelope fluctuations of flattened maskers are minimized, but not completely eliminated as would be the case for pure-tone stimuli. Crest factors in decibels are displayed next to their corresponding temporal envelope power spectra and were 10.7 and 4.4 dB for fluctuating and flattened maskers, respectively. Temporal envelope power and crest factor calculations displayed in Fig. 1 were obtained by averaging the temporal envelope power and crest factors of 2000 independently generated noise waveforms. Waveforms in Fig. 1 were obtained by passing maskers through a 1- ERB_N wide gammatone filter centered on the probe frequency. The amplitude of these waveforms was scaled to best illustrate the temporal envelope fluctuations of the maskers.

Stimuli were generated using MATLAB®-based (The MathWorks, Natick, MA) custom software (Bidelman et al. 2015) and output to the right earphone of EARTONE-5A (3 M, Minneapolis, MN) insert earphones after being routed through a LynxTWO-B (Lynx Studio Technology, Costa Mesa, CA) sound card (sampling rate, 44.1 kHz; 24-bit resolution) and a headphone buffer (Tucker-Davis Technologies, HB7, Alachua, FL).

Exp. 1 evaluated the effects of masker temporal envelope fluctuation on masker level at threshold in temporal effect conditions. Independent variables were masker temporal envelope fluctuation (fluctu-

Fig. 1. Schematic representation of the stimuli (upper panel) and masker temporal envelope power spectra (lower panel), along with corresponding waveform insets. The 4000-Hz, 6-ms probe was presented either near the onset or temporal center of the 400-ms masker (only the temporal center condition is shown). The masker bandwidth was 0.5 equivalent rectangular bandwidths (228 Hz) and was centered linearly on the probe frequency. Crest factors were calculated at the output of an auditory filter centered on the probe frequency in order to quantify the degree of masker fluctuation for fluctuating (fluc., diamonds) and flattened (flat, squares) maskers

ating and flattened), probe delay (onset and temporal center), and probe level. Several probe levels were used because the temporal effect may be missed in some subjects if tested at only one level (Bacon 1990). The dependent variable was masker level at threshold. Exp. 2 evaluated the effect of temporal envelope fluctuations on probe level at threshold for fluctuations occurring before, or simultaneous with, the presentation of the probe. This was done by presenting either a flattened or fluctuating precursor prior to the onset of either flattened or fluctuating maskers.

Procedure

Subjects sat in a double-walled, sound-attenuating room while participating in the experiments. Masker level at threshold (Exp. 1) was obtained using a three-interval, three-alternative forced choice task, with a

two-up, one-down adaptive rule to converge on 70.7 % correct detection (Levitt 1971). For Exp. 2, probe level at threshold was measured (one-up, two-down adaptive rule, 70.7 % correct detection). The step size of the adaptive track was 5 dB for the first four reversals and 2 dB for the remaining eight reversals. Threshold of the adaptive track was defined as the average level of the final eight reversals. Observation intervals were marked by lights and separated by 500 ms. Two intervals contained the masker; the other contained the masker and probe. Subjects pushed a button on a keyboard to indicate in which interval the probe was perceived and visual feedback was used to indicate a correct or incorrect response. Thresholds were first measured for probes in quiet and were 19, 19, 20, 25, 23, 20, 19, 27, 23, and 17 dB SPL for S1–S10, respectively. Subjects were then familiarized with the experiments by measuring masker level at threshold for two repetitions of a 60-dB SPL probe presented near the onset and temporal center of fluctuating maskers. During phase I of training, listeners practiced detecting 45- and 60-dB SPL probes presented near the temporal center of fluctuating and flattened maskers (4 conditions $\times$ 1 threshold per condition = 4 threshold estimates). Phase II of training was identical to phase I except that the probe was presented near the onset of the masker. Thresholds measured during familiarization and training were discarded.

During data collection for Exp. 1, conditions were randomized by probe delay, probe level, and masker fluctuation. For Exp. 2, conditions were randomized by probe delay, masker fluctuation, and precursor fluctuation. Thresholds from three (Exp. 1) or two (Exp. 2) adaptive tracks were obtained for each condition and averaged to compute the final threshold. If the standard deviation of these adaptive tracks exceeded 5 dB, an additional adaptive track was obtained and included in the average.¹ An additional adaptive track was obtained four times: twice in one subject and twice in another subject in Exp. 1 and once each in two subjects in Exp. 2. The first threshold measured during a given data collection session served as a “warm-up” and was discarded.

EXPERIMENT 1

For this experiment, masker level at threshold was measured for 45-, 60-, and 75-dB SPL probes presented near the onset and temporal center of fluctuating

and flattened maskers to determine the effects of masker fluctuation on the temporal effect.

Results

Figure 2 displays results from Exp. 1. On average, SNR (probe level–masker level) at threshold was constant or decreased slightly as a function of probe level for probes presented near the onset of fluctuating (Fig. 2a, open diamonds) and flattened (Fig. 2a, open squares) maskers, respectively. For probes presented near the temporal center of fluctuating maskers (Fig. 2a, solid diamonds), SNR at threshold increased and then decreased slightly as a function of probe level and was always greater than SNRs at threshold for probes presented near the onset of fluctuating maskers. These findings resulted in a negative temporal effect (Fig. 2b, diamonds), which is consistent with previous studies that used 4000-Hz probes and on-frequency NBN maskers with bandwidths ≤ 10 % of the probe frequency (Bacon and Smith 1991; Bacon and Liu 2000). The opposite trend occurred for probes presented near the temporal center of flattened maskers (Fig. 2a, solid squares), where SNR at threshold decreased as a function of probe level and was always lower than SNRs at threshold for probes presented near the onset of flattened maskers. These findings resulted in a positive temporal effect that increased with increasing probe level, up to ~ 7 dB (Fig. 2b, squares), consistent with the temporal effect observed for on-frequency tonal maskers, which have flat envelopes (Bacon and Viemeister 1985). The temporal effect obtained with fluctuating and flattened maskers was submitted to a repeated-measures analysis of variance (rmANOVA) with probe level (45, 60, 75 dB SPL) and masker type (fluctuating, flattened) as repeated measures. The main effect of masker type [$F(1,5) = 29.26$, $p < 0.005$] was significant, while the main effect of probe level [$F(2,10) = 1.65$, $p = 0.24$] and the probe level $\times$ masker type interaction [$F(2,10) = 3.51$, $p = 0.07$] were insignificant. These analyses show that the temporal effect was larger for flattened, compared to fluctuating maskers, regardless of probe level.

Figure 2c displays the difference between SNRs at threshold for probes presented with fluctuating and flattened maskers and located near the onset (dashed line, open circles) or temporal center (solid line, closed circles) of maskers. This “flat-fluctuating difference” (dB) was obtained by subtracting masker level at threshold for the fluctuating masker from thresholds measured with the flattened masker. When averaged across probe level, the mean (standard error) flat-fluctuating differences were 3.7 dB (0.6 dB) and 11.1 dB (1.1 dB) for the probe onset and

¹ For Exp. 1, subject S2 was unable to return to complete a fourth adaptive track on one condition with a standard deviation greater than 5 dB; thus, data for this subject are the average of the original three adaptive tracks.

Fig. 2. Results from Exp. 1. **a** Average signal-to-noise ratio (SNR; probe level–masker level) at threshold as a function of probe level and probe delay for fluctuating (“fluc,” diamonds) and flattened (“flat,” squares) maskers. SNRs at threshold for probes presented at the masker’s onset (2 ms) and the masker’s temporal center (197 ms) are plotted as open and filled symbols, respectively. **b** The temporal effect, as calculated by subtracting masker levels at threshold for probes presented at the masker’s onset from those obtained for probes presented at the masker’s temporal center. **c** The flat-fluctuating difference (see text) for probes presented at the masker’s onset (open circles, dashed lines) or at the masker’s temporal center (closed circles, solid lines). In **b** and **c**, the horizontal dashed gray lines are plotted at 0 dB for reference. Error bars are one standard error of the mean

probe temporal-center conditions, respectively. The flat-fluctuating difference was submitted to an rmANOVA with probe level (45, 60, 75 dB SPL) and probe delay (2, 197 ms) as repeated measures. The main effects of probe level [$F(2,10) = 11.9$, $p < 0.005$] and probe delay [$F(1,5) = 29.26$, $p < 0.005$] were

significant, while the probe level $\times$ probe delay interaction [$F(2,10) = 3.51$, $p = 0.07$] was insignificant. These analyses show that the flat-fluctuating difference increases with increasing probe level and is strongest for probes presented near the temporal center of maskers, compared to probes presented at the masker’s onset.

EXPERIMENT II

Data from Exp. 1 revealed greater flat-fluctuating differences for probes presented near the temporal center, compared to the onset of maskers (Fig. 2c). In addition, fluctuating maskers were more effective when probes were presented near the temporal center, compared to the onset of these maskers (Fig. 2a, diamonds). These findings are consistent with a relationship between detection thresholds for short probes and the AC-coupled envelope power (ENV_{AC}) of the masker in an analysis window temporally centered on the probe. Specifically, detection thresholds are expected to worsen as masker ENV_{AC} increases in this window. Masker ENV_{AC} is smallest for probes presented near the temporal center of flattened maskers and increases when (1) the masker’s temporal envelope is made to fluctuate, and (2) the probe is presented near the onset of the masker, due to the masker’s ramp. The hypothesis that detectability of a short tone depends on masker ENV_{AC} is consistent with that of Svec et al. (2016), who showed greater threshold elevation for probes preceded by forward maskers with fluctuating, compared to flattened, temporal envelopes. The relationship between detection thresholds for short probes and masker ENV_{AC} was evaluated by presenting either silence or flattened or fluctuating precursors immediately prior to either flattened or fluctuating maskers. The precursor had the same spectrum as the masker, and the masker and precursor were presented at 75 dB SPL. Probe level was adapted to estimate thresholds. The masker’s onset ramp overlapped the precursor’s offset ramp by 2 ms to avoid an audible gap between precursor and masker.²

² A 2-ms overlap between precursor and masker ramps eliminated audible gaps/transients for all conditions except for flattened precursors followed by flattened maskers. A rapid shift in the fine structure occurring at the junction of the low-noise precursor and low-noise noise masker sometimes resulted in an audible transient. To rectify this, a long low-noise noise stimulus with a duration equal to the combined duration of the precursor and masker was generated instead of generating separate precursors and maskers. This solution was possible because the levels and spectra of the precursors and maskers were equivalent.

Results

Figure 3 displays SNR at threshold for probes presented near the onset (left group of bars) or temporal center (right group of bars) of fluctuating (light gray bars) or flattened (dark gray bars) maskers. Maskers were preceded by silence (silent precursor) or by fluctuating or flattened precursors, as indicated by temporal envelope waveforms below the x-axis. An rmANOVA with precursor type (silent, fluctuating, flattened), masker type (fluctuating, flattened), and probe delay (2, 197 ms) as repeated measures revealed significant main effects of precursor type ($F[2,8] = 25.3$, $p < 0.001$) and masker type ($F[1,4] = 152.5$, $p < 0.001$), as well as significant precursor type $\times$ masker type ($F[2,8] = 56.13$, $p < 0.001$), precursor type $\times$ probe delay ($F[2,8] = 34.1$, $p < 0.001$), and masker type $\times$ probe delay ($F[1,4] = 46.9$, $p < 0.01$) interactions.

Pairwise comparisons conducted on the main effect of precursor type revealed higher SNRs at thresholds for fluctuating, compared to silent ($t[4] = -4.4$, $p < 0.05$), or flattened ($t[4] = -7.2$, $p < 0.01$) precursors. Similarly, pairwise comparisons conducted on the main effect of masker type revealed significantly higher SNRs at threshold for fluctuating compared to flattened maskers ($t[4] = -12.4$, $p < 0.001$). Post hoc testing for the precursor type $\times$ masker type interaction revealed that flattening the masker's temporal envelope results in a significantly smaller decrease in SNR at threshold for fluctuating precursors, compared to silent ($t[4] = -8.8$, $p < 0.005$), or flattened

($t[4] = -9.2$, $p < 0.005$), precursors, regardless of probe delay. Post hoc testing for the precursor type $\times$ probe delay interaction revealed that the effect of changing the precursor type from fluctuating to flattened resulted in significantly larger decreases in SNR at threshold for probes delayed by 2 ms, compared to 197 ms from the masker's onset ($t[4] = -8.6$, $p < 0.005$). Finally, post hoc testing for the masker type $\times$ probe delay interaction revealed that flattening the masker's temporal envelope resulted in significantly greater improvements in probe threshold for probes presented near the temporal center, compared to the onset of maskers ($t[4] = -6.8$, $p < 0.005$), consistent with the results from Exp. 1.

A correlation analysis was performed to test the hypothesis that probe thresholds are associated with ENV_{AC} of precursor/masker pairs in an analysis window temporally centered on the probe. To obtain ENV_{AC} , stimulus waveforms were half-wave rectified, low-pass filtered at 150 Hz, and squared to convert the temporal envelope into power-like units. ENV_{AC} was calculated as the standard deviation of the squared temporal envelope within a rectangular analysis window temporally centered on the probe. Due to the stochastic nature of noise maskers and precursors, 100 stimuli were generated and processed, after which ENV_{AC} was calculated on the average of these processed stimuli. Several analysis window durations were tested including 12, 25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000 ms. Coefficients of determination (R^2) for the correlation between aver-

FIG. 3. Results from Exp. 2. SNR at threshold is plotted for either fluctuating (light gray bars) or flattened (dark gray bars) maskers preceded by silence, or by fluctuating or flattened precursors (see temporal envelope waveforms below the x-axis). The group of bars to the left is for probes presented near the onset of the masker. The group of bars to the right is for probes presented near the temporal center of the masker

age SNR at threshold and ENV_{AC} are shown in Table 1 for each analysis window. R^2 was greater than 0.90 for analysis windows between 100 and 500 ms and declined to as low as 0.75 for the shortest and longest analysis windows. A plot of average SNR at threshold as a function of ENV_{AC} for the best-fitting analysis window (400 ms) is shown in Fig. 4, along with the associated regression line. Symbols in Fig. 4 represent different pairs of precursor/masker temporal envelopes, as indicated by the temporal envelope waveforms in the legend. Open and filled symbols are for probes presented near the onset and temporal center of the masker, respectively. When the correlation analysis between SNR at threshold and ENV_{AC} (400 ms analysis window) was applied to thresholds from individual subjects, rather than the average thresholds, R^2 was greater than 0.90 for all subjects, except S2, where R^2 was 0.71. These results suggest that the strong relationship between SNR at threshold and ENV_{AC} observed for the mean data also holds for data from individual subjects.

DISCUSSION

This study shows that, for normal-hearing adults, the lack of a temporal effect for NBN maskers (bandwidth $\leq 10\%$ of probe frequency) centered on the probe frequency is partially due to the inherent temporal envelope fluctuations of these maskers. This conclusion is supported by the finding that a positive temporal effect was observed for NBN maskers with flattened, but not fluctuating, temporal envelopes (Fig. 2b). Moreover, quantifying temporal envelope fluctuations as ENV_{AC} accounts for 94 % of the

variance in SNRs at threshold for probes presented in the context of fluctuating or flattened NBN maskers and precursors (Fig. 4). The analysis window durations most effective at accounting for this variance (50–400 ms) are consistent with time constants of other auditory temporal processes such as temporal integration (Gerken et al. 1990) and forward masking (Moore et al. 1988).

The strong relationship between ENV_{AC} and probe thresholds is consistent with the success of models with temporal envelope-based decision variables in accounting for perceptual performance in simultaneous masking (Richards 1992), amplitude modulation detection (Ewert and Dau 2000), and speech perception (Jorgensen and Dau 2011). Moreover, interpreting detection thresholds for short probes in longer simultaneous maskers in terms of ENV_{AC} is consistent with the finding that thresholds for short probes are higher when presented near the onset *and* offset (Elliott 1965) of longer maskers, compared to thresholds for probes presented near the temporal center of maskers. A parsimonious interpretation of the temporal effect observed for flattened NBN maskers (Figs. 2b and 4, open and closed left-pointing triangles) is that listeners rely on fluctuations of the stimulus temporal envelope to detect the probe, and that the onset ramp of the masker interferes with listeners' ability to segregate the probe

TABLE 1	
Coefficients of determination (R^2) observed for the correlation between SNRs at threshold and the AC-coupled temporal envelope power of the precursor/masker stimuli in an analysis window temporally centered on the probe. Results for the best fitting analysis window are italicized	
<i>Analysis window (ms)</i>	R^2
12	0.75
25	0.84
50	0.89
100	0.91
200	0.91
300	0.91
400	<i>0.94</i>
500	0.90
600	0.88
700	0.83
800	0.79
900	0.77
1000	0.75

Fig. 4. The results of a correlation analysis between SNRs at threshold and the average AC-coupled temporal envelope power in a 400-ms, rectangular window temporally centered on the probe. Open and filled symbols represent probes presented near the onset and temporal center of the masker, respectively. Temporal envelope waveforms in the legend indicate conditions with fluctuating or flattened maskers preceded by silence, or preceded by fluctuating or flattened precursors

from the masker. The hypothesis that temporal envelope fluctuations may lead to stimulus uncertainty has been proposed and supported by previous studies using fluctuating and flattened maskers in simultaneous (Buss et al. 2006) and forward masking (Svec et al. 2015, 2016). An advantage of interpreting the data from Exps. I and II in terms of ENV_{AC} is the ability to account for the temporal effect and for the effects of fluctuating and flattened precursors under a single unified framework. Despite this, such an interpretation deviates from the mechanisms commonly evoked to explain the temporal effect, such as a reduction in cochlear amplifier gain over the time course of the masker (Schmidt and Zwicker 1991; von Klitzing and Kohlrausch 1994; Strickland 2001; Bacon and Savel 2004; Jennings et al. 2011).

The temporal effect observed with flattened NBN maskers in this study is qualitatively different from the temporal effect observed in previous studies involving BBN maskers, which have been well explained by models based on cochlear gain reduction (Strickland 2001; Jennings et al. 2016). This qualitative difference is illustrated in Fig. 5a, which compares SNR at threshold as a function of probe level for broadband Gaussian noise maskers (Jennings et al. 2016) and narrowband low-fluctuating noise maskers (current study) for probes presented near the onset (open symbols) and temporal center (closed symbols) of maskers. SNR at threshold was computed at the output of a 1-ERB_N wide filter centered on the probe frequency (4000 Hz for both studies). The temporal effect for BBN maskers increases with increasing probe level. This increase is due to increasing SNR at threshold with increasing probe level for probes presented near the onset of the masker (Fig. 5a, open triangles), while SNR at threshold remains constant with increasing level for probes presented near the temporal center of the masker (Fig. 5a, closed triangles). Higher SNRs at threshold for moderate-to-high level probes, compared to low-level probes, in the probe onset condition has been interpreted and successfully modeled as evidence for greater cochlear compression at moderate-to-high, compared to low probe levels (Strickland 2001; Jennings et al. 2016). In other words, the internal representation of the probe+masker interval grows slowly (due to compression), while that of the lower-level masker-only interval grows more quickly, resulting in shallow growth of masking (Fig. 5b, open triangles). This interpretation is supported by the finding that SNR at threshold for probes presented at the masker's onset is better (lower) than normal in subjects with temporary (Champlin and McFadden 1989; McFadden and Champlin 1990) or permanent cochlear hearing loss (Bacon and Takahashi 1992; Strickland and Krishnan 2005; Jennings et al. 2016), who are expected to have

Fig. 5. Comparison of temporal effects measured with broadband Gaussian noise maskers (triangles, Jennings et al. 2016), and narrowband low-fluctuating maskers (squares, current study) in terms of **a** SNR at threshold and **b** growth of masking. SNR and masker level at threshold for probes presented near the masker's onset and the masker's temporal center are displayed as open and closed symbols, respectively. Dotted lines provide a reference for linear (1 dB/dB) growth of masking

reduced cochlear compression. Constant SNRs at threshold as a function of probe level in the probe temporal center condition have been interpreted as evidence for a reduction in cochlear compression over the time course of the masker for normal-hearing listeners (von Klitzing and Kohlrausch 1994). A reduction in gain is expected to linearize cochlear responses, leading to growth of masking with a unity slope (Fig. 5b, closed triangles).

Similar to the temporal effect with BBN maskers (e.g., Jennings et al. 2016), the temporal effect with narrowband flattened maskers (current study) increases with increasing probe level. Despite this, the relationship between SNR at threshold and probe level is difficult to account for with the gain reduction

model, which assumes that the probe is detected through an auditory filter centered on the probe frequency, and that the temporal effect is explained by a decrease in cochlear compression over the time course of the masker (Strickland 2001). Based on these assumptions, the constant SNRs at threshold as a function of probe level in the probe onset condition (Fig. 5a, open squares) and the growth of masking with a unity slope (Fig. 5b, open squares) suggest that the internal representations of the probe+masker and masker-only intervals grow at the same rate. This similar rate of growth indicates that the responses of the probe+masker and masker-only intervals through an auditory filter centered on the probe frequency grow linearly, or are subjected to the same degree of cochlear compression (Nelson and Schroder 1997). Thus, unlike results from BBN maskers, there is no evidence to support or refute the idea that cochlear compression accounts for the relatively higher SNRs at threshold for probes presented near the onset, compared to the temporal center of NBN maskers. The decrease in SNR at threshold as a function of probe level (Fig. 5a, closed squares) and the steep growth of masking (Fig. 5b, closed squares) for probes presented near the temporal center of flattened NBN maskers are difficult to explain in terms of a reduction in cochlear gain. The growth of masking slope greater than unity implies that the internal representation of the probe+masker interval grows more quickly than that of the masker-only interval. A reduction in gain is expected to linearize the cochlear response through the auditory filter centered on the probe frequency. Thus, growth of masking with a unity slope is expected, unless the output of the probe+masker interval exceeds the compression breakpoint. If this were the case, shallow growth of masking is expected, rather than the steep growth of masking observed in Fig. 5b (closed squares).

Although ENV_{AC} accounts for the temporal effects observed in this study, it is not expected to account for the temporal effect observed with BBN maskers. For example, if the probe is detected through an auditory filter centered on the probe frequency (i.e., critical band model), ENV_{AC} at the output of such a filter is expected to be large in response to BBN due to the filter's accentuation of slow temporal envelope fluctuations (Dau et al. 1999). Thus, ENV_{AC} predicts that results for BBN maskers will be similar to the results for fluctuating NBN maskers: higher (poorer) thresholds for probes presented near the temporal center, compared to the onset of the masker. This prediction is inconsistent with the 10–25 dB *improvement* in probe thresholds observed with BBN maskers as the probe is delayed from the onset, to the temporal center of the masker (Bacon 1990).

ENV_{AC} may partially explain the temporal effect with BBN noise maskers if the critical band model is discarded in favor of a multiband model (Gilkey and Robinson 1986; Davidson et al. 2006). In this case, a decision variable based on the summed response across several auditory filters would have the effect of flattening the internal representation of the masker's temporal envelope. This reduction in the effective ENV_{AC} predicts a 5–10-dB temporal effect, similar to that observed for flattened NBN maskers (Fig. 2b, squares); however, this falls short of the 10–25-dB temporal effect observed for BBN maskers. The notion that listeners combine information across several auditory filters to improve performance has been well established in studies on comodulation masking release (Hall et al. 1984) and profile analysis (Green 1987). Additionally, a multiband model may account for the potential effects of off-frequency listening through cochlear filters basal to the probe frequency, where response growth is expected to be linear (Moore et al. 1999). Such off-frequency listening effects are not expected for BBN maskers and may partially account for differences in growth of masking slopes among NBN and BBN maskers.

Temporal effects for BBN and NBN maskers differ in their SNRs at threshold (Fig. 5a), growth of masking slopes (Fig. 5b), and their ability to be accounted for by the gain reduction model and/or by ENV_{AC} . These findings are consistent with the general phenomenon of the temporal effect originating from several different underlying mechanisms, where the contribution of each mechanism depends on stimulus settings (e.g., masker bandwidth). Similar conclusions about multiple mechanisms contributing to the temporal effect were offered by Carlyon and White (1992) and Bacon and Savel (2004). The significance of the current study is the finding that ENV_{AC} contributes to the temporal effects measured with NBN maskers and may contribute to temporal effects observed with other maskers. The dependence of the temporal effect on factors such as masker spectrum (Bacon and Smith 1991), probe frequency (Strickland 2001), probe duration (Zwicker 1965a), hearing impairment (Bacon and Takahashi 1992), and masker temporal envelope characteristics (current study) highlights the complexity of the temporal effect and motivates the need to interpret such data in the context of an auditory model. Based on the results of the current study, future efforts to model the temporal effect should simulate the effects of ENV_{AC} in addition to simulating other potential mechanisms of the temporal effect including (among others) compression and cochlear gain reduction.

SUMMARY AND CONCLUSIONS

This study shows that temporal effects measured with NBN maskers depend on the inherent fluctuations of the masker's temporal envelope. Detection thresholds for short probes presented with fluctuating and flattened precursors and maskers are consistent with a decision variable based on the AC-coupled power of the precursor and masker temporal envelopes. Temporal effects measured with NBN maskers and BBN maskers differ in terms of SNRs at threshold and growth of masking, which is consistent with these temporal effects originating from separate mechanisms. Success of future efforts to model the temporal effect may depend on the use of a decision variable based on the AC-coupled power of the temporal envelope.

ACKNOWLEDGMENTS

The authors would like to thank Judy Dubno, Jayne Ahlstrom, Donal Sinex, Bill Yost, Sarah Yoho-Leopold, Ali Almishaal, and Jessica Chen for helpful discussions and comments about this work. Comments from Peggy Nelson and Elizabeth Strickland were helpful in improving the quality of this manuscript.

FUNDING INFORMATION

This research was completed while the first author was a mentee in a mentoring program for clinician scientists in hearing and balance disorders sponsored by the National Institutes of Health (U24 DC012079). This work was supported in part by Grant No. K23 DC014752 from the National Institutes of Health/National Institute on Deafness and Other Communication Disorders (PI: Jennings).

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of Interest The authors declare that they have no conflict of interest.

REFERENCES

ALMISHAAL A, BIDELMAN GM, JENNINGS SG (2017) Notched-noise precursors improve detection of low-frequency amplitude modulation. *J Acoust Soc Am* 141:324–333

BACON SP (1990) Effect of masker level on overshoot. *J Acoust Soc Am* 88:698–702

BACON SP, VIEMEISTER NF (1985) The temporal course of simultaneous tone-on-tone masking. *J Acoust Soc Am* 78:1231–1235

BACON SP, SMITH MA (1991) Spectral, intensive, and temporal factors influencing overshoot. *Q J Exp Psychol A* 43:373–399

BACON SP, TAKAHASHI GA (1992) Overshoot in normal-hearing and hearing-impaired subjects. *J Acoust Soc Am* 91:2865–2871

BACON SP, LIU L (2000) Effects of ipsilateral and contralateral precursors on overshoot. *J Acoust Soc Am* 108:1811–1818

BACON SP, SABEL S (2004) Temporal effects in simultaneous masking with on- and off-frequency noise maskers: effects of signal frequency and masker level. *J Acoust Soc Am* 115:1674–1683

BIDELMAN GM, JENNINGS SG, STRICKLAND EA (2015) *PsyAcoustX*: a flexible MATLAB® package for psychoacoustics research. *Front Psychol* 6(1498):1–11

BUSS E, HALL JW 3RD, GROSE JH (2006) Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *J Acoust Soc Am* 120:2777–2788

CARLYON RP, WHITE LJ (1992) Effect of signal frequency and masker level on the frequency regions responsible for the overshoot effect. *J Acoust Soc Am* 91:1034–1041

CHAMPLIN CA, MCFADDEN D (1989) Reductions in overshoot following intense sound exposures. *J Acoust Soc Am* 85:2005–2011

DAU T, VERHEY J, KOHLRAUSCH A (1999) Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers. *J Acoust Soc Am* 106:2752–2760

DAVIDSON SA, GILKEY RH, COLBURN HS, CARNEY LH (2006) Binaural detection with narrowband and wideband reproducible noise maskers. III. Monaural and diotic detection and model results. *J Acoust Soc Am* 119:2258–2275

DAVIS H (1983) An active process in cochlear mechanics. *Hear Res* 9:79–90

ELLIOTT LL (1965) Changes in the simultaneous masked threshold of brief tones. *J Acoust Soc Am* 38:738–746

EWERT SD, DAU T (2000) Characterizing frequency selectivity for envelope fluctuations. *J Acoust Soc Am* 108:1181–1196

FLETCHER H (1940) Auditory patterns. *Rev Mod Phys* 12:47–65

GERKEN GM, BHAT VK, HUTCHISON-CLUTTER M (1990) Auditory temporal integration and the power function model. *J Acoust Soc Am* 88:767–778

GILKEY RH, ROBINSON DE (1986) Models of auditory masking: a molecular psychophysical approach. *J Acoust Soc Am* 79:1499–1510

GLASBERG BR, MOORE BC (1990) Derivation of auditory filter shapes from notched-noise data. *Hear Res* 47:103–138

GREEN DM (1987) *Profile analysis: auditory intensity discrimination*. Oxford University Press, Oxford

GUINAN JJ (2010) Cochlear efferent innervation and function. *Curr Opin Otolaryngol Head Neck Surg* 18:447–453

HALL JW, HAGGARD MP, FERNANDES MA (1984) Detection in noise by spectro-temporal pattern analysis. *J Acoust Soc Am* 76:50–56

JENNINGS SG, STRICKLAND EA (2010) The frequency selectivity of gain reduction masking: analysis using two equally effective maskers. In: Lopez-Poveda EA, Palmer AR, Meddis R (eds) *Advances in auditory physiology, psychophysics and models*. Springer, New York, pp 47–58

JENNINGS SG, STRICKLAND EA (2012) Evaluating the effects of olivocochlear feedback on psychophysical measures of frequency selectivity. *J Acoust Soc Am* 132:2483–2496

JENNINGS SG, STRICKLAND EA, HEINZ MG (2009) Precursor effects on behavioral estimates of frequency selectivity and gain in forward masking. *J Acoust Soc Am* 125:2172–2181

JENNINGS SG, HEINZ MG, STRICKLAND EA (2011) Evaluating adaptation and olivocochlear efferent feedback as potential explanations of psychophysical overshoot. *J Assoc Res Otolaryngol* 12:345–360

JENNINGS SG, AHLSTROM JB, DUBNO JR (2016) Effects of age and hearing loss on overshoot. *J Acoust Soc Am* 140:2481–2493

JORGENSEN S, DAU T (2011) Predicting speech intelligibility based on the signal-to-noise envelope power ratio after modulation-frequency selective processing. *J Acoust Soc Am* 130:1475–1487

KOHLRAUSCH A, FASSEL R, VAN DER HEIJDEN M, KORTEKAAS R, VAN DE PAR S, OXENHAM AJ, PÜSCHEL D (1997) Detection of tones in low-noise noise: further evidence for the role of envelope fluctuations. *Acust Acta Acust* 83:659–669

- KRULL V, STRICKLAND EA (2008) The effect of a precursor on growth of forward masking. *J Acoust Soc Am* 123:4352–4357
- LEVITT H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477
- McFADDEN D (1989) Spectral differences in the ability of temporal gaps to reset the mechanisms underlying overshoot. *J Acoust Soc Am* 85:254–261
- McFADDEN D, CHAMPLIN CA (1990) Reductions in overshoot during aspirin use. *J Acoust Soc Am* 87:2634–2642
- MOORE BC, GLASBERG BR, PLACK CJ, BISWAS AK (1988) The shape of the ear's temporal window. *J Acoust Soc Am* 83:1102–1116
- MOORE BC, VICKERS DA, PLACK CJ, OXENHAM AJ (1999) Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *J Acoust Soc Am* 106:2761–2778
- NELSON DA, SCHRODER AC (1997) Linearized response growth inferred from growth-of-masking slopes in ears with cochlear hearing loss. *J Acoust Soc Am* 101:2186–2201
- RICHARDS VM (1992) The detectability of a tone added to narrow bands of equal-energy noise. *J Acoust Soc Am* 91:3424–3435
- ROVERUD E, STRICKLAND EA (2010) The time course of cochlear gain reduction measured using a more efficient psychophysical technique. *J Acoust Soc Am* 128:1203–1214
- ROVERUD E, STRICKLAND EA (2015) The effects of ipsilateral, contralateral, and bilateral broadband noise on the mid-level hump in intensity discrimination. *J Acoust Soc Am* 138:3245–3261
- SCHMIDT S, ZWICKER E (1991) The effect of masker spectral asymmetry on overshoot in simultaneous masking. *J Acoust Soc Am* 89:1324–1330
- STRICKLAND EA (2001) The relationship between frequency selectivity and overshoot. *J Acoust Soc Am* 109:2062–2073
- STRICKLAND EA (2004) The temporal effect with notched-noise maskers: analysis in terms of input-output functions. *J Acoust Soc Am* 115:2234–2245
- STRICKLAND EA (2008) The relationship between precursor level and the temporal effect. *J Acoust Soc Am* 123:946–954
- STRICKLAND EA, KRISHNAN LA (2005) The temporal effect in listeners with mild to moderate cochlear hearing impairment. *J Acoust Soc Am* 118:3211–3217
- SVEC A, DUBNO JR, NELSON PB (2015) Effects of inherent envelope fluctuations in forward maskers for listeners with normal and impaired hearing. *J Acoust Soc Am* 137:1336–1343
- SVEC A, DUBNO JR, NELSON PB (2016) Inherent envelope fluctuations in forward maskers: effects of masker-probe delay for listeners with normal and impaired hearing. *J Acoust Soc Am* 139:1195–1203
- VIEMEISTER NF (1979) Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am* 66:1364–1380
- VON KLITZING R, KOHLRAUSCH A (1994) Effect of masker level on overshoot in running- and frozen-noise maskers. *J Acoust Soc Am* 95:2192–2201
- WRIGHT BA (1997) Detectability of simultaneously masked signals as a function of masker bandwidth and configuration for different signal delays. *J Acoust Soc Am* 101:420–429
- ZWICKER E (1965A) Temporal effects in simultaneous masking by white-noise bursts. *J Acoust Soc Am* 37:653–663
- ZWICKER E (1965B) Temporal effects in simultaneous masking and loudness. *J Acoust Soc Am* 38:132–141