A further test of the linearity of temporal summation in forward masking (L)

Christopher J. Plack[a] and Samuele Carcagno
Department of Psychology, Lancaster University, Lancaster, LA1 4YF, United Kingdom

Andrew J. Oxenham
Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, Minnesota 55455

(Received 29 January 2007; revised 10 May 2007; accepted 2 August 2007)

An experiment tested the hypothesis that the masking effects of two nonoverlapping forward maskers are summed linearly over time. First, the levels of individual noise maskers required to mask a brief 4-kHz signal presented at 10-, 20-, 30-, or 40-dB sensation level (SL) were found. The hypothesis predicts that a combination of the first masker presented at the level required to mask the 10-dB SL signal and the second masker presented at the level required to mask the 20-dB SL signal, should produce the same amount of masking as the converse situation (i.e., the first masker presented at the level required to mask the 20-dB SL signal and the second masker presented at the level required to mask the 10-dB SL signal), and similarly for the 30- and 40-dB SL signals. The results were consistent with the predictions. © *2007 Acoustical Society of America.*

[DOI: 10.1121/1.2775287]

PACS number(s): 43.66.Dc, 43.66.Mk, 43.66.Ba, 43.64.Kc [JHG] Pages: 1880–1883

I. INTRODUCTION

Forward masking refers to the decrease in the detectability of a signal as a result of prior stimulation by a masker. Forward masking has been shown to be highly nonlinear in listeners with normal hearing. For example, a given increase in the level of a forward masker produces a much smaller increase in the signal level at threshold, for signal levels less than about 30 dB SPL (Jesteadt *et al.*, 1982; Moore and Glasberg, 1983; Munson and Gardner, 1950). However, it has been demonstrated that these nonlinear effects can be simulated by a model that incorporates a compressive nonlinearity, representing the response of the basilar membrane (BM), prior to a *linear* leaky integrator or temporal window (Plack and Oxenham, 1998; Plack *et al.*, 2002). The output of the temporal window is a linear weighted sum over time of a quantity proportional to the square of BM velocity.

One of the predictions of linear summation is that the contribution of a stimulus to masking should be unaffected by the presence of stimuli before and/or after; the internal representations of two stimuli separated in time should be *independent*. In a recent study the prediction was tested using two nonoverlapping forward maskers (Plack *et al.*, 2006). It was demonstrated that the contribution of the second masker to masking was unaffected by the first masker, even when the first masker in the sequence rendered the second masker inaudible. The present experiment is a further test of this prediction. Consider a masking situation with two consecutive forward maskers, M1 and M2. The levels of the forward maskers are chosen so that in the first condition M1 has the level required to mask a signal with level L_s dB and M2 has the level required to mask a signal with level L_s+x dB, and in the second condition M1 has the level required to mask a signal with level L_s+x dB and M2 has the level required to mask a signal with level L_s dB. We assume that for a given signal level the masking effect (in terms of the temporal window model, this is the output of the window in response to the masker at the time of signal presentation) is constant at threshold, irrespective of the time of presentation or duration of the masker, for example. Hence, the combined effect of the maskers is the same in both conditions: In each case the total masking effect is the sum of that required to mask a signal with level L_s dB and that required to mask a signal with level L_s+x dB. Because the masking effect is the same, the linear-summation hypothesis predicts that the signal thresholds in the two conditions should be *identical*. If, however, there is a nonlinear interaction between the maskers, for example if the first masker reduces the effectiveness of the second when the first masker is higher in level, the prediction will not hold. The present study therefore provides an empirical test of this prediction, which arises from the hypothesis of linear summation, to further test the validity of the temporal window model as an account of auditory temporal masking.

II. METHOD

A. Stimuli

The signal was a 4000-Hz pure tone. The first masker in the sequence (M1) was a Gaussian noise, bandpass filtered between 2800 and 5600 Hz (3-dB cutoffs, 90 dB/octave). The second masker (M2) was a Gaussian noise, bandpass filtered between 3400 and 4800 Hz. The signal had a total duration of 4 ms, which consisted of 2-ms raised-cosine onset and offset ramps (no steady state). Quoted levels are peak equivalent sound pressure levels. M1 had a total duration of 200 ms, including 2-ms onset and offset ramps and a 196-ms

---

[a] Author to whom correspondence should be addressed. Electronic mail: c.plack@lancaster.ac.uk

0001-4966/2007/122(4)/1880/4/$23.00

TABLE I. The absolute thresholds for the signal and the results of Phase 1 of the experiment. The masker levels at threshold are given in decibel spectrum level for each sensation level of the signal. Standard errors are given in parentheses.

	Absolute threshold for signal (dB SPL)	10 dB SL		20 dB SL		30 dB SL		40 dB SL	
		M1	M2	M1	M2	M1	M2	M1	M2
L1	21.9 (0.5)	−2.3 (0.5)	7.6 (0.9)	13.3 (1.5)	24.4 (2.0)	28.9 (1.4)	34.3 (1.5)	47.3 (1.1)	44.0 (0.9)
L2	12.9 (0.4)	−11.0 (0.3)	14.8 (1.1)	5.1 (1.6)	31.8 (0.6)	19.4 (1.1)	40.5 (0.4)	30.0 (0.5)	47.8 (0.8)
L3	11.1 (1.1)	−14.9 (0.8)	8.9 (4.0)	0.3 (1.1)	25.8 (1.1)	15.3 (1.7)	34.3 (2.0)	25.0 (0.7)	39.2 (0.7)
L4	14.3 (0.4)	−16.1 (0.4)	−6.3 (2.2)	−6.7 (0.7)	11.9 (2.9)	−1.9 (3.0)	24.5 (3.3)	11.7 (0.9)	36.8 (3.0)
Mean	15.0 (2.4)	−11.1 (3.1)	6.3 (4.5)	3.0 (4.2)	23.5 (4.2)	15.4 (6.4)	33.4 (3.3)	28.5 (7.4)	41.9 (2.5)

steady-state portion. M2 had a total duration of 6 ms, including 2-ms onset and offset ramps and a 2-ms steady-state portion. The end of M1 coincided with the start of M2. The silent interval between the end of M2 and the start of the signal was 4 ms.

The experiment was controlled by custom-made software from a PC workstation located outside a double-walled sound-attenuating booth. All stimuli were generated digitally with 32-bit resolution and were output by an RME Digi96/8 PAD 24-bit soundcard set at a clocking rate of 48 kHz. The headphone output of the soundcard was fed via a patch panel in the sound booth wall to Sennheiser HD580 headphones without filtering or amplification. Stimuli were presented to the right ear. Each listener sat in the booth and decisions were recorded via a computer keyboard. Listeners viewed a computer monitor through a window in the sound booth. Lights on the monitor display flashed on and off concurrently with each stimulus presentation and provided feedback at the end of each trial.

B. Procedure

Four normally hearing listeners (ages 23–29) took part in the experiments. Listeners were given at least 2 h training on the conditions before data collection. The procedure was similar to that described in previous articles (Plack and O'Hanlon, 2003; Plack *et al.*, 2006). For both absolute and masked thresholds, a three-interval forced-choice procedure was used with an interstimulus interval of 300 ms. Threshold was determined using a two-up one-down (Phase 1) or a two-down one-up (absolute thresholds and Phase 2) adaptive procedure that tracked the 70.7 percent correct point on the psychometric function (Levitt, 1971). The step size was 4 dB for the first four turnpoints, and was reduced to 2 dB for 12 subsequent turnpoints. The mean of the last 12 turnpoints was taken as the threshold estimate for each block of trials. At least four estimates were made for each condition and the results averaged.

First, absolute thresholds for the signal were measured for each listener. The main experiment was then conducted in two phases:

(i) Phase 1. The level of the signal was fixed at 10-, 20-, 30-, or 40-dB sensation level (SL). In each case, the levels of M1 and M2 required to mask the signal when presented individually were determined. For M1, these maskers are designated M1(10), M1(20), M1(30), and M1(40) and similarly for M2.[1]

(ii) Phase 2. The threshold for the signal was determined in the presence of M1, M2, and M1+M2 (M1 and M2 combined). M1 and M2 were presented at the levels determined in Phase 1. The combined masker conditions were M1(10)+M2(20), M1(20)+M2(10), M1(30)+M2(40), M1(40)+M2(30).

III. RESULTS AND ANALYSIS

A. Results

The absolute thresholds for the signal, and the results of Phase 1, are presented in Table I. Despite the similarity in absolute thresholds, there is considerable variability in the masker levels required to mask the signal, particularly at the higher signal levels. However, this is not uncommon in situations in which the signal and/or masker are at sound levels greater than about 30 dB SPL and are therefore subject to strong compression (e.g. Plack and Drga, 2003; Plack and O'Hanlon, 2003).

The results of Phase 2 are presented in Fig. 1. The thresholds for the single masker conditions should be roughly equal to the signal levels presented in Phase 1. Although this is generally the case at low levels, the mean data show that the thresholds are less than expected at the higher levels (30 and 40 dB SL correspond to mean signal levels of 45 and 55 dB SPL; see Table I). The combined masker thresholds are, in general, considerably higher than the highest single masker threshold. With the exceptions of listeners L1 and L4 in the M1(20)/M2(10) condition, "excess" masking is observed in all cases. This is indicative of a compressive system (Oxenham and Moore, 1995). A repeated-measures analysis of variance (ANOVA) [overall level (10/20, 30/40) × masker sequence (low M1/high M2, high M1/low M2)] conducted on the combined masker thresholds revealed a highly significant effect of overall level [$F(1,3)=1090$, $p<0.0005$] but no effect of masker sequence [$F(1,3)=0.17$, $p=0.71$] and no interaction [$F(1,3)=2.24$, $p=0.23$]. The results from L1 and L4 in the M1(20)/M2(10) condition were somewhat anomalous, in that less excess

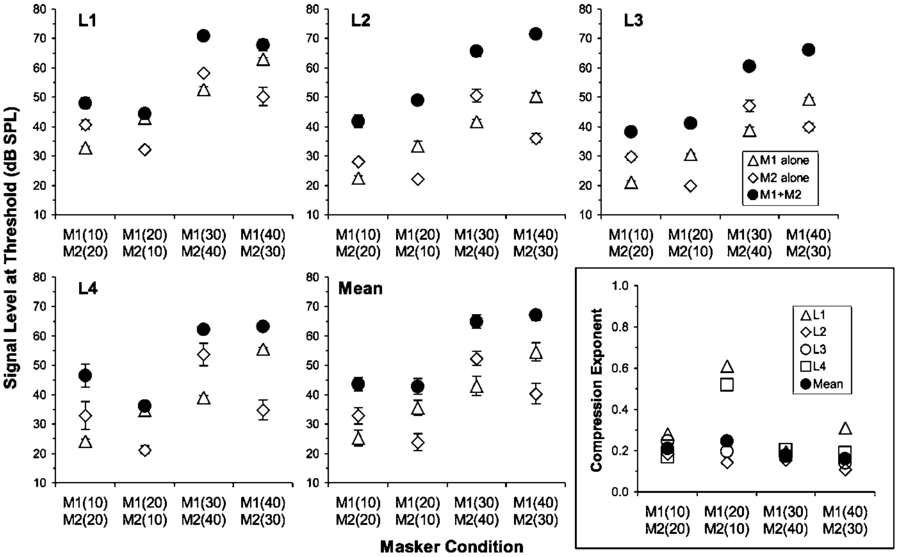

FIG. 1. The results of Phase 2 of the experiment, showing signal thresholds in the presence of the single (open symbols) and combined (closed symbols) maskers. Error bars show standard errors. The lower right panel shows compression exponents derived from the data (see the text for details).

masking was observed, resulting in an apparent difference between that condition and the M1(10)/M2(20) condition. This could be interpreted as evidence that the higher-level M1 was affecting the influence of the lower-level M2 for these two listeners. However, in the same subjects at the higher level, and in the other two listeners at both levels, no such evidence was found. Overall, as indicated by the statistical analysis, there was no evidence that the combined masker thresholds are affected by the order of presentation of the maskers.

B. Derivation of compression exponents

The technique for deriving compression exponents from the signal threshold data was the same as that used by Plack and O'Hanlon (2003). It was assumed that the ratio of the internal (i.e., postcochlear) signal magnitude to the internal (or effective) masker magnitude is a constant at signal threshold. It was also assumed that the internal signal magnitude is a power-law transformation of physical signal intensity. Hence,

$$I_{\mathrm{M}} = kS_{\mathrm{M}}^{c}, \tag{1}$$

where I_{M} represents the internal effect of the masker, S_{M} is the physical signal intensity at masked threshold, c is the compression exponent, and k is a constant. It was assumed further that the effect of combining two maskers is a linear summation of their individual effects. Hence,

$$I_{\mathrm{M1+M2}} = I_{\mathrm{M1}} + I_{\mathrm{M2}}. \tag{2}$$

Substituting Eq. (1) in Eq. (2), and factoring out the constant k, leaves

$$S_{\mathrm{M1+M2}}^{c} = S_{\mathrm{M1}}^{c} + S_{\mathrm{M2}}^{c}. \tag{3}$$

If $S_{\mathrm{M1+M2}}$ (the signal intensity at threshold for the combined maskers), S_{M1} (the signal threshold for M1 alone), and S_{M2} (the signal threshold for M2 alone) are all known, it is possible to determine the compression exponent c. This was achieved using the SOLVER algorithm in Microsoft Excel.

The derived compression exponents are shown in Fig. 1. The exponents for the mean data were derived from the mean signal thresholds, rather than being simply the mean of the compression exponents across the four listeners. The values for the higher-level stimuli, 0.18 for M1(30)/M2(40) and 0.16 for M1(40)/M2(30), are close to the values reported in previous studies (e.g., Oxenham and Plack, 1997; Plack and O'Hanlon, 2003). Also as expected, there is a tendency for the values to increase at low levels, particularly for listeners L1 and L4. A repeated-measures ANOVA [overall level (10/20, 30/40) × masker sequence (low M1 / high M2, high M1 / low M2)] conducted on the compression exponents revealed a nearly significant effect of overall level [$F(1,3)=9.08$, $p=0.057$] no effect of masker sequence [$F(1,3)=1.28$, $p=0.34$] and no interaction [$F(1,3)=2.66$, $p=0.20$].

IV. CONCLUSIONS

Overall, the results are consistent with the hypothesis that the masking effects of nonoverlapping forward maskers combine in a linear manner (although this is only tested for masker and signal levels up to about 70 dB SPL in the present study). Furthermore, the compression exponents derived from the signal thresholds are consistent with previous psychophysical and physiological (Ruggero *et al.*, 1997; Yates *et al.*, 1990) measures of BM compression, suggesting that the main nonlinearity in forward masking is the result of cochlear processing. After quasi-instantaneous cochlear compression, the effects of combining nonoverlapping maskers can be well described by a time-invariant linear system.

ACKNOWLEDGMENTS

The authors thank the Editor and two anonymous reviewers for constructive comments on an earlier version of the manuscript. The work was supported by EPSRC (UK) Grant No. GR/N07219, BBSRC (UK) Grant No. BB/D012953/1, and NIDCD Grant No. R01 DC 03909.

---

[1] Phase 1 provided the masker levels required to the test the hypothesis using the combined masker thresholds from Phase 2 (Sec. III A). Although the second analysis in terms of compression exponents (Sec. III B) could be done using signal thresholds measured with arbitrary masker levels in

Phase 2, Phase 1 is necessary to at least approximately match the effectiveness of the maskers. If one masker is much more effective than the other it will dominate masking in the combined case, and the derivation of the exponent will be unreliable.

Jesteadt, W., Bacon, S. P., and Lehman, J. R. (**1982**). "Forward masking as a function of frequency, masker level, and signal delay," J. Acoust. Soc. Am. **71**, 950–962.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Moore, B. C. J., and Glasberg, B. R. (**1983**). "Growth of forward masking for sinusoidal and noise maskers as a function of signal delay: Implications for suppression in noise," J. Acoust. Soc. Am. **73**, 1249–1259.

Munson, W. A., and Gardner, M. B. (**1950**). "Loudness patterns—A new approach," J. Acoust. Soc. Am. **22**, 177–190.

Oxenham, A. J., and Moore, B. C. J. (**1995**). "Additivity of masking in normally hearing and hearing-impaired subjects," J. Acoust. Soc. Am. **98**, 1921–1934.

Oxenham, A. J., and Plack, C. J. (**1997**). "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **101**, 3666–3675.

Plack, C. J., and Drga, V. (**2003**). "Psychophysical evidence for auditory compression at low characteristic frequencies," J. Acoust. Soc. Am. **113**, 1574–1586.

Plack, C. J., and O'Hanlon, C. G. (**2003**). "Forward masking additivity and auditory compression at low and high frequencies," J. Assoc. Res. Otolaryngol. **4**, 405–415.

Plack, C. J., and Oxenham, A. J. (**1998**). "Basilar-membrane nonlinearity and the growth of forward masking," J. Acoust. Soc. Am. **103**, 1598–1608.

Plack, C. J., Oxenham, A. J., and Drga, V. (**2002**). "Linear and nonlinear processes in temporal masking," Acta. Acust. Acust. **88**, 348–358.

Plack, C. J., Oxenham, A. J., and Drga, V. (**2006**). "Masking by inaudible sounds and the linearity of temporal summation," J. Neurosci. **26**, 8767–8773.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). "Basilar-membrane responses to tones at the base of the chinchilla cochlea," J. Acoust. Soc. Am. **101**, 2151–2163.

Yates, G. K., Winter, I. M., and Robertson, D. (**1990**). "Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range," Hear. Res. **45**, 203–220.