Method-of-adjustment measures of informational masking between auditory streams

Stanley Sheft

Parmly Hearing Institute, Loyola University Chicago, 6525 North Sheridan Road, Chicago, Illinois 60626
ssheft@luc.edu

William A. Yost

Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287
William.Yost@asu.edu

Abstract: A method-of-adjustment procedure was used to measure thresholds for detecting a continuous sequence of brief 2-kHz tonal pulses in the presence of random-frequency masking sequences. Masker pulses consisted of either one or eight sinusoidal components and were either synchronous or asynchronous with the signal pulses. Effects of pulse rate and asynchronous gating were generally consistent with a reduction in informational masking due to segregation of the signal and masker streams. Despite use of continuous stimulus presentation to encourage stream segregation, masking was still obtained from most listeners in most conditions.
© 2008 Acoustical Society of America

1. Introduction

Informational masking refers to difficulty in either signal detection or discrimination beyond that predicted by energetic masking between a signal and masker (for a review, see Kidd *et al.*, 2008). Though initial studies emphasized masker uncertainty as the basis of the interference, more recent considerations have also attributed a role to similarity between the signal and masker (e.g., Durlach *et al.*, 2003; Watson, 2005). Involvement of similarity allows for possible reduction in informational masking by enhancing distinction between the signal and masker. Working with auditory sequences, Kidd *et al.* (1994) demonstrated a release from informational masking through stimulus manipulations intended to promote perceptual segregation of the signal and masker auditory streams. Subsequent work by Durlach *et al.* (2003) confirmed this result with studies by Kidd *et al.* (2003) and Micheyl *et al.* (2007) supporting the interpretation linking the masking release to stream segregation.

As with informational masking, models of auditory streaming invoke involvement of stimulus similarity in the perceptual effect (McNally and Handel, 1977; Moore and Gockel, 2002). For streaming, stimulus-event similarity is seen as the basis of stream coherence, whereas for informational masking, similarity lessens component resolution; with separate streams defined as signal and masker, local similarity encourages stream segregation to enhance signal detection. A common basis for the two effects suggests strong potential for release from informational masking as a result of stream segregation. Though presenting multiple stimulus pulses, overall stimulus duration was relatively brief (roughly 0.5–2.4 s) in the informational-masking studies cited previously. Conversely, studies of the buildup of stream segregation indicate that the effect accumulates over longer time spans (e.g., Anstis and Saida, 1985), especially if attention is not refocused or diverted (Cusack *et al.*, 2004). The present study evaluated signal detection in the presence of random-frequency masker streams, utilizing continuous stimulus presentation intended to fully promote signal stream segregation from a masker of undetermined sequential coherence. Along with overall duration, gating characteristics affect segregation, both in studies of streaming (e.g., Dannenbring and Bregman, 1978; Turgeon *et al.*, 2005) and informational masking (Kidd *et al.*, 1994; Neff, 1991, 1995; Durlach *et al.*, 2003; Leibold
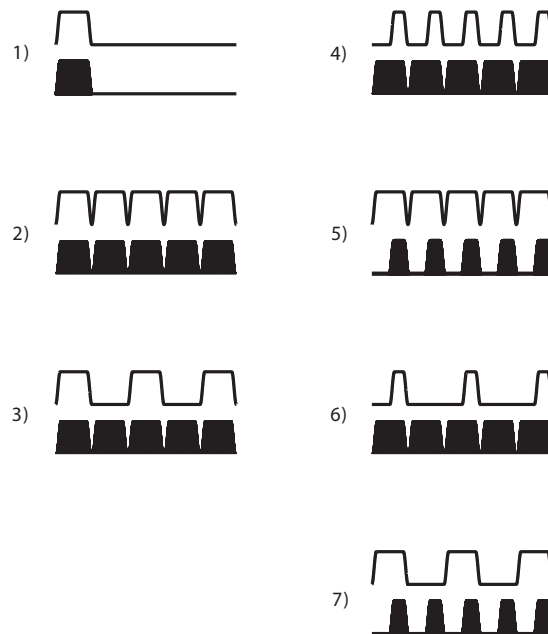
Fig. 1. Schematic illustration of the seven masking conditions. For each condition, the signal stream is illustrated at the top with the open trace and the masker stream below with the filled trace. Signal- and masker-pulse duration was either 40 or 80 ms with concurrent pulses either synchronously or asynchronously gated (see the text for additional details).

and Neff, 2007; Micheyl *et al.*, 2007). Consequently, experimental conditions evaluated the effect of temporal disparities in terms of both asynchronous gating of concurrent signal and masker pulses along with presentation of only alternate signal pulses against the masker stream.

2. Method

Thresholds for detecting a continuous stream of 2-kHz signal pulses were measured for 16 listeners both in quiet and in the presence of a continuous masker stream. Five configurations for the signal sequence were paired with three masker-sequence configurations for a total of seven masking conditions (see Fig. 1). Within the signal and masker sequences, pulse duration was either 40 or 80 ms. When the signal and masker pulses shared a common 80-ms duration, signal and masker pulses were either synchronously gated with an interstimulus interval (ISI) of 420 ms (condition 1) or 0 ms (condition 2), or every other pulse was omitted from the signal stream with the masker ISI maintained at 0 ms (condition 3). In the remaining conditions, signal- and masker-pulse durations differed with the streams aligned so that either signal-pulse onset was delayed by 40 ms relative concurrent masker pulses (conditions 4 and 6), or masker-pulse onset was delayed (conditions 5 and 7). As in condition 3, the signal was presented only during alternate masker pulses in conditions 6 and 7. All signal and masker pulses were shaped with 10-ms cosinusoidal rise/fall times.

Masker pulses consisted of either one or eight sinusoidal components. With a single masker component, masker frequency was either fixed at 1417 Hz, or for each pulse was drawn at random from a uniform distribution evenly spaced on a logarithmic scale ranging from 500 to 7000 Hz, excluding a "protected" region of 1600–2500 Hz about the signal frequency. The component phase was randomly selected from a uniform distribution between 0 and 2π. A similar scheme for random selection was used for each component of the eight-component maskers; in addition, relative component amplitude was randomly drawn from a Rayleigh distribution. With both the one- and eight-component maskers, overall masker level was fixed at 65 dB SPL. Dell PCs with 24-bit Echo Gina 3G soundcards were used for stimulus generation
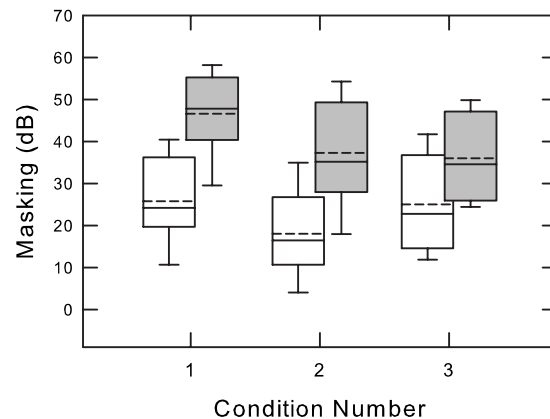
Fig. 2. As a box plot, group masking data from conditions 1–3 in which the common signal- and masker-pulse duration was 80 ms. Masker frequencies were randomly selected with the open boxes for results obtained with the single-component masker, and the gray boxes for the eight-component-masker conditions. Box boundaries indicate the 25th and 75th percentiles, with error bars extending to the 10th and 90th percentiles. Group median and mean threshold values are indicated within each box by horizontal solid and dashed lines, respectively.

and experimental control. Following analog conversion at a 22.05-kHz sampling rate, stimuli were low-pass filtered at 8 kHz and presented diotically through Sennheiser HD 280 Pro headphones with the listeners seated in a double-walled soundproof booth.

To accommodate continuous stimulus presentation, thresholds were measured with a method-of-adjustment procedure. Both Watson *et al.* (2002) and Yost *et al.* (2007) used a method-of-adjustment procedure in investigations of the effects of masker uncertainty, with Watson and his co-workers reporting that in some cases the procedure allowed for listeners to more rapidly adopt a detection strategy that lessened the extent of informational masking. Once a listener initiated a threshold run in the present study, signal level was continuously controlled by the listener via adjustment of a slider on a graphical user interface displayed on a computer monitor. Listeners were instructed to adjust the level so that they could just detect the signal stream. Listeners were encouraged to adopt a bracketing strategy to set threshold levels, and also to allow time between level adjustments for stream formation. In the masking conditions, the procedure began with a 5-s presentation of the signal stream in isolation at 85 dB SPL as a cue. Following a 1.5-s silent pause, the signal and masker streams were presented concurrently with listeners then given control of signal level. In all conditions, a single method-of-adjustment threshold was collected from each of the 16 listeners across two listening sessions. Experimental protocol was approved by the Institutional Review Board of Loyola University Chicago.

3. Results and discussion

As in most studies of informational masking, individual differences characterized results from all conditions. To convey both the distribution of threshold values and data trends, results are displayed as box plots. Figure 2 shows results from conditions 1–3 in which the common signal- and masker-pulse duration was 80 ms. Masking in dB is the threshold difference obtained with and without the masker present. In all conditions, masker frequencies were randomly selected with the open boxes representing conditions with single-component masker pulses, and the gray boxes indicating results obtained with eight-component maskers. Across conditions 1–3, masking was greater in the multi- than single-component conditions. This effect of the number of masker components is consistent with past results from informational-masking studies (e.g., Neff, 1991, 1995; Kidd *et al.*, 1994). With either the single- and eight-component maskers, masking diminished with decreasing ISI from 420 to 0 ms between conditions 1 and 2. In that stream coherence is not expected with the longer ISI (e.g., Bregman *et al.* 2000), a masking
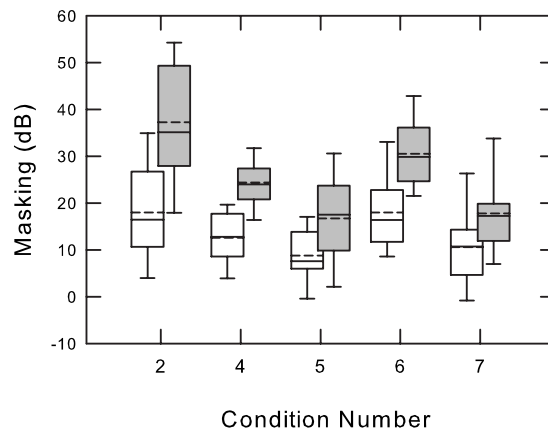
Fig. 3. Same as Fig. 2 for conditions 4–7 in which either the signal- or masker-pulse onset was delayed 40 ms. For comparison, results from condition 2 are replotted from Fig. 2.

release due to segregation of the constant-frequency signal stream from the random-frequency maskers is anticipated only with the 0-ms ISI. An inverse relationship between masking and ISI was a central result used by Kidd *et al.* (2003) to support interpretation of stream segregation as a potential factor in release from informational masking.

In condition 3, presentation of only alternate signal pulses created a temporal disparity between the signal and masker streams. Kidd *et al.* (1994) reported a reduction in masking with signal presentation only during alternate pulses of identical masker samples. The masking release was obtained despite the drop in signal energy due to alternation. With the omission of every other signal pulse and frequency randomization across masker pulses within constrained ranges, Micheyl *et al.* (2007) found that the extent of masking was consistent with change in signal energy. That is, the temporal disparity generated by omitting signal pulses did not aid performance, and in fact led to threshold elevation. In the present conditions in which masker frequency was randomly varied, comparison of results from conditions 2 and 3 indicates that signal alternation either had no effect or led to a slight increase in masking.

These observations were supported by a two-factor repeated-measures analysis of variance (ANOVA), which showed significant main effects of number of masker components $[F(1,15)=59.0; p<0.001]$ and condition $[F(2,30)=5.4; p=0.01]$, along with a significant interaction $[F(2,30)=9.3; p=0.001]$. Tukey's post-hoc pairwise comparisons confirmed a significant effect of ISI, and that only with a single-component masker did omitting signal pulses through alternation significantly affect performance.

Figure 3 shows results from conditions 4–7 in which either the signal- or masker-pulse onset was delayed 40 ms relative to the other. Results from condition 2 in which the signal and masker pulses were synchronously gated are shown on the left for comparison. In all conditions, masker frequencies were randomly selected with box shading again distinguishing the number of masker components. As in the initial data set, increasing the number of masker components from one to eight increased the amount of masking. Comparing group threshold values in conditions 2, 4, and 5 indicates that delay of either the signal or masker onset reduced the amount of masking with a greater masking release observed with delay of the masker rather than the signal. As with condition 3 of the first data set, the stimulus configurations of conditions 6 and 7 presented only alternate signal pulses against the continuous masker stream. Similar to the results from the first data set, temporal disparity due alternation of signal pulses on average had little to no effect.

A two-factor repeated-measures ANOVA supported these observations, showing significant main effects of number of masker components $[F(1,15)=105; p<0.001]$ and condition $[F(4,60)=29.1; p<0.001]$, along with a significant interaction $[F(4,60)=6.7; p<0.001]$.
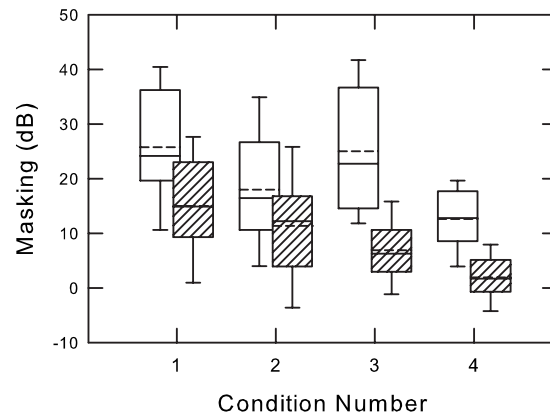
Fig. 4. Same as Fig. 2 for conditions 1–4 with a single-component masker. Right-hatched boxes represent results obtained with the masker frequency fixed at 1417 Hz; the open boxes replot data from Figs. 2 and 3 in which masker frequency was randomized across pulses.

Tukey's post-hoc pairwise comparisons confirmed a significant effect of delaying either the signal- or masker-pulse onset, except in condition 4 with a single-component masker and a 40-ms signal delay. Comparison of performance in conditions 4 and 6 and conditions 5 and 7 indicated a significant effect of omitting signal pulses only with the eight-component masker and the 40-ms signal delay. Among the four reference conditions, this comparison involves the one for which thresholds were, on average, the highest, suggesting that the absence of a release from masking was not due to a ceiling effect.

  In conditions with a single 200-ms masker burst, Neff (1991, 1995), and Durlach *et al.* (2003) reported substantial release from informational masking when signal duration was reduced so that it was no longer synchronously gated with the masker. On average, smaller reductions in masking were obtained in the present work with delay of the signal-pulse onset (condition 4), especially in the presence of the single-component masker stream. Differences in stimulus configuration are most likely are the major reason for the distinction (see Sheft, 2008, for a review of influence of configuration on effect of gating asynchrony). In the current study, masker duration was 80 ms to encourage stream segregation with continuous presentation, and signal delay was then limited to 40 ms, half the pulse duration. Though the 40-ms delay is shorter than used in the previous informational-masking studies, results from the streaming studies of Dannenbring and Bregman (1978) and Turgeon *et al.* (2005) demonstrate significant segregation with delay in this range, which in the latter study often approached baseline performance measures. In conditions in which performance may be based on detection of asynchrony rather than signal segregation per se, listeners are sensitive to delays as brief as 1 ms despite masker uncertainty (Huang and Richards, 2006).

  In the present work, restriction on masking release due to asynchrony may relate to the suggestion of Dannenbring and Bregman (1978), and one supported by results from Turgeon *et al.* (2005), that in the context of streaming, asynchrony is effective when it leads to presentation of a segment of the signal or target without overlapping masker presentation, a configuration atypical of other informational-masking studies. Although the current results show that this speculation is not exclusively true, group thresholds were lowest in condition 5 in which the signal-pulse onset preceded masker onset by 40 ms. Presumably the relatively small threshold elevations that persist in condition 5 reflect a conventional backward-masking effect.

  In the introduction, two factors were discussed as contributing to informational masking, uncertainty and similarity. The final data set compares performance in conditions in which the frequency of the single-component masker was either fixed at 1417 Hz or randomized across pulses. The fixed-masker conditions allow for evaluation of similarity without uncertainty. In Fig. 4, results from conditions 1–4 with the 1417-Hz masker are shown with right-

hatched boxes; thresholds obtained with randomizing masker frequency are replotted from Figs. 2 and 3 with open boxes. In all cases, group thresholds were lower when the masker frequency was fixed rather than randomized across pulses. Two-tailed matched-sample t tests indicated that each difference was statistically significant ($t = 2.5, 2.4, 5.8$, and 7.9; $p < 0.03$ for conditions 1–4). The trend in which the amount of masking progressively dropped across conditions 1–4 indicates that streaming, augmented by cross-stream temporal disparities, can reduce masking in conditions without frequency uncertainty. Similar results in the context of informational masking have recently been reported by Leibold and Neff (2007) and Buss (2008).

Overall, significant results in the present work, with the number of individual exceptions from the total in parentheses, are as follows: Effects of the number of masker components (2 of 112), pulse ISI (6 of 32), omitting signal pulses (8 of 32), asynchronous gating of the signal and masker (4 of 48), and frequency randomization with a single masker component (10 of 64). In two cases, consideration of exceptions calls for further comment. For omission of signal pulses, the effect itself was obtained with only half of the configurations in which it was used. Regarding the effect of frequency randomization, 6 of the 10 exceptions were from condition 2. As noted earlier, stream segregation tends to become more efficient with introduction of cross-stream temporal disparities as was done in conditions 3 and 4 for which the total number of exceptions dropped to only three.

The motivation of this study was to evaluate the effects of auditory streaming on informational masking. A general criticism of investigations of informational masking is that the effect is often discussed in terms of what it is not, that is, something other than energetic masking. Garner (1974) presented a framework for considering information processing in the context of the perception of structure, arguing that, functionally, information and structure are identical terms. Informational masking can then be viewed as a limitation on information transmission due to poorly defined or insufficient structure. Auditory scene analysis represents perception of structure. By enhancing structure, auditory stream segregation may increase the potential for information transmission with this seen experimentally as a release from informational masking.

Acknowledgments

References and links

Anstis, S., and Saida, S. (**1985**). "Adaptation to auditory streaming of frequency-modulated tones," J. Exp. Psychol. Hum. Percept. Perform. **11**, 257–271.

Bregman, A. S., Ahad, P. A., Crum, P. A. C., and O'Reilly, J. (**2000**). "Effects of time intervals and tone durations on auditory stream segregation," Percept. Psychophys. **62**, 626–636.

Buss, E. (**2008**). "The effect of masker level uncertainty on intensity discrimination," J. Acoust. Soc. Am. **123**, 254–264.

Cusack, R., Deeks, J., Aikman, G., and Carlyon, R. P. (**2004**). "Effects of location, frequency region, and time course of selective attention on auditory scene analysis," J. Exp. Psychol. Hum. Percept. Perform. **30**, 643–656.

Dannenbring, G. L., and Bregman, A. S. (**1978**). "Streaming vs. fusion of sinusoidal components of complex tones," Percept. Psychophys. **24**, 369–376.

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G. (**2003**). "Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity," J. Acoust. Soc. Am. **114**, 368–379.

Garner, W. R. (**1974**). *The Processing of Information and Structure* (Lawrence Erlbaum Assoc., Potomac, Maryland).

Huang, R., and Richards, V. M. (**2006**). "Coherence detection: Effects of frequency, frequency uncertainty, and onset/offset delays," J. Acoust. Soc. Am. **119**, 2298–2304.

Kidd, G., Mason, C. R., and Deliwala, P. S. (**1994**). "Reducing informational masking by sound segregation," J. Acoust. Soc. Am. **95**, 3475–3480.

Kidd, G., Mason, C. R., and Richards, V. M. (**2003**). "Multiple bursts, multiple looks, and stream coherence in the release from informational masking," J. Acoust. Soc. Am. **114**, 2835–2845.

Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. I. (**2008**). "Informational masking," in *Auditory Perception of Sound Sources*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer Science+Business, New York).

Leibold, L. J., and Neff, D. L. (**2007**). "Effects of masker-spectral variability and masker fringes in children and adults," J. Acoust. Soc. Am. **121**, 3666–3676.

Micheyl, C., Shamma, S. A., and Oxenham, A. (**2007**). "Hearing out repeating elements in randomly varying multitone sequences: a case of streaming?" in *Hearing—From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer, Berlin).

McNally, K. A., and Handel, S. (**1977**). "Effect of element composition on streaming and the ordering of repeating sequences," J. Exp. Psychol. **3**, 451–460.

Moore, B. C. J., and Gockel, H. (**2002**). "Factors influencing sequential stream segregation," Acta. Acust. Acust. **88**, 320–332.

Neff, D. L. (**1991**). "Forward masking by maskers of uncertain frequency content," J. Acoust. Soc. Am. **89**, 1314–1323.

Neff, D. L. (**1995**). "Signal properties that reduce masking by simultaneous, random-frequency maskers," J. Acoust. Soc. Am. **98**, 1909–1920.

Sheft, S. (**2008**). "Envelope processing and sound-source perception," in *Auditory Perception of Sound Sources*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer Science+Business, New York).

Turgeon, M., Bregman, A. S., and Roberts, B. (**2005**). "Rhythmic masking release: Effects of asynchrony, temporal overlap, harmonic relations, and source separation on cross-spectral grouping," J. Exp. Psychol. Hum. Percept. Perform. **31**, 939–953.

Watson, C. S. (**2005**). "Some comments on informational masking," Acta. Acust. Acust. **91**, 502–512.

Watson, C. S., Kidd, G. R., and Pok, S. V. (**2002**). "Attentional focus and the method of adjustment revisited," J. Acoust. Soc. Am. **112**, 2243.

Yost, W. A., Cabel, K., and Sheft, S. (**2007**). "Using method of adjustment to measure informational masking using multitone maskers," J. Acoust. Soc. Am. **121**, 3132.