Informational masking increases the costs of monitoring multiple channels

Gerald Kidd, Jr.

*Department of Speech, Language and Hearing Sciences and Hearing Research Center, Boston University, Boston, Massachusetts 02215
gkidd@bu.edu*

Virginia M. Richards

*Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104
richards@psych.upenn.edu*

Christine R. Mason

*Department of Speech, Language and Hearing Sciences and Hearing Research Center, Boston University, Boston, Massachusetts 02215
cmason@bu.edu*

Frederick J. Gallun

*National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, Oregon 97239
frederick.gallun@va.gov*

Rong Huang

*Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104
rongh@sas.upenn.edu*

Abstract: This study examined the costs of simultaneously monitoring two frequency regions. Listeners detected low- and high-frequency tones in a 2I4AFC procedure. On every trial, each signal was presented in either the first or second interval independently. Comparison of thresholds in single- and dual-signal conditions provided an estimate of the costs. Thresholds were obtained in quiet, in notched-filtered noise, and in randomized multitone maskers. No cost was found in quiet, whereas large costs were found for the masked conditions, especially for the multitone masker. These results suggest that costs of dividing attention in frequency depend on both signal and non-signal channels.

© 2008 Acoustical Society of America

PACS numbers: 43.66.Dc, 43.66.Ba, 43.66.Lj [QJF]

Date Received: May 5, 2008 **Date Accepted:** July 17, 2008

1. Introduction

If the task of a listener is to detect the presence of a pure-tone signal in a background of noise, uncertainty about the frequency of the signal elevates detection threshold relative to the threshold that is found when the frequency of the signal is known. However, the elevation in threshold due to signal frequency uncertainty is relatively small, only about 3 dB for a wide range of possible frequencies (e.g., [Green, 1961](#)). In fact, attempting to explain the small effect of signal frequency uncertainty was one of the early problems addressed by researchers applying signal detection theory to auditory psychophysics (e.g., [Tanner *et al.*, 1956](#); [Veniar, 1958](#); [Creelman, 1960](#); [Green, 1961](#); [Sorkin *et al.*, 1973](#)). The basic problem was that the variance of the decision statistic computed from monitoring multiple frequency channels simultaneously, or from shifting the center frequency of a single tunable filter across a wide range of frequencies, appeared to be too great to account for the empirical results. Later work ([Buus *et al.*, 1986](#)) demonstrated that the apparent discrepancy was less than originally considered. However, taken as a whole,

the conclusion that is drawn from all of these studies is that the *cost*¹ of monitoring multiple frequency channels is relatively low.

The various attempts at modeling the effects of signal frequency uncertainty have in common the underlying assumption that the costs associated with monitoring different frequency regions arise solely from the channels containing, or potentially containing, signal energy. This assumption makes intuitive sense because frequency channels that never include the signal contain no information useful in solving the task and, as such, an ideal observer would not consider them in forming a decision rule. Human listeners, however, often use decision strategies that are less than ideal. The phenomenon of informational masking, in which masker energy remote in frequency from the signal energy exerts a significant effect on performance, is one such example. [Neff and Green \(1987\)](#) were the first to demonstrate that a masker composed of randomly drawn sinusoids having a different composition on every presentation produced large amounts of masking that could not be attributed to peripheral overlap of excitation. Since that initial report many investigators have used the multitone masking paradigm to examine central influences in auditory masking (see the review by [Kidd *et al.*, 2008](#)).

Although it is clear that large amounts of informational masking may be produced by highly uncertain multitone maskers, it is not known whether there is an interaction between informational masking and the costs of signal frequency uncertainty. Such an interaction seems plausible because the decrease in performance associated with signal frequency uncertainty, and the informational masking caused by masker uncertainty and/or similarity, both presumably reflect limitations on higher-level processes. This is in contrast to energetic masking, which is assumed to be due primarily to overlap of excitation in the auditory periphery. In the present study, we attempt to address this issue by assessing the costs of signal frequency uncertainty under three levels of masker uncertainty: no masker (quiet, no masker uncertainty), Gaussian noise masker (low masker uncertainty; intended to produce primarily energetic masking), and randomized multitone masker varying in composition on every stimulus presentation (high uncertainty; intended to produce large amounts of informational masking).

2. Methods

2.1 Listeners

A total of six young adult college students served as paid participants in these experiments. The subjects had normal hearing as determined by routine audiometry.

2.2 Stimuli

The stimuli were computer-generated tone and noise bursts that were played through a 16 bit Digital-to-Analog Converter (DAC) and programmable analog attenuators, mixers, low-pass filters (7.5 kHz), and headphone buffers (Tucker-Davis Technologies, Alachua, FL).

The frequencies of the signal tones were 558 and 1791 Hz. The duration of the masker was 205 ms. In all cases except one the signal was also 205 ms and presented synchronously with the masker. In the one exception described in the following, the onset of a 195 ms signal was delayed by 10 ms relative to that of the masker. All stimuli were gated by applying 5 ms cosine-squared ramps and the durations specified are total durations including ramps and steady-state portions.

There were two types of maskers: double-notched-filtered Gaussian noise and random-frequency multitone complexes. The Gaussian noise masker was generated digitally from 200 to 5000 Hz and had spectral notches centered geometrically on the two signal frequencies each spanning a width of 28% of the center frequency. The noise was first generated in the frequency domain and then inverse Fourier transformed producing extremely steep filter slopes. The level of the noise was held constant at 50 dB SPL (about 14 dB/Hz within the passbands). The random-frequency multitone maskers were drawn from the same frequency range as the noise bands with “protected regions” (i.e., in which masker components were not allowed) equal in width to the notches in the noise masker. The number of masker tones varied across intervals such that the total number of tones—signal tone(s) plus masker tones—was

constant at eight. Thus, for example, if both signal tones were in interval 1, six randomly chosen masker tones were presented in interval 1, while eight randomly chosen masker tones were presented in interval 2. If there was one signal tone in each interval, seven masker tones were presented in each interval. The purpose was to restrict judgments based purely on the number of tones present. There was always a different random draw of masker frequencies on each interval. The level per masker component was held constant at 41 dB SPL. There were two presentation conditions for the multitone masker. In one case, the 205 ms signal and masker were gated on and off together and that condition is referred to as “synchronous.” The other condition, in which the 195 ms signal was delayed by 10 ms from the 205 ms multitone masker onset, is designated “asynchronous.” The asynchronous condition was intended to be intermediate between the notched-noise and the synchronous multitone masker (i.e., the asynchrony was intended to release some of the expected informational masking).

2.3 Procedures

Listening was monaural with stimuli presented to the subject’s preferred ear through a calibrated TDH-50 earphone. The listeners were tested individually in double-walled sound attenuating booths.

Detection thresholds were obtained using a 3-down 1-up adaptive procedure that estimates the 79.8% correct point on the psychometric function (Levitt, 1971). The main variables in the experiment were whether the listener’s task was to detect one or two signals on each trial and the presence/type of masker that was used. In dual-signal conditions the listener made two separate detection judgments, one corresponding to each signal and two separate tracks provided a threshold estimate for each. The four alternatives (signal frequency by presentation interval) were represented on the response keys so only a single response was required to indicate the listener’s choice. For all listeners, masked thresholds were obtained for both single-signal and dual-signal conditions in quiet (no masker) and in the presence of double-notched-filtered Gaussian noise and random-frequency multitone maskers. A minimum of 15 blocks of trials was obtained in each condition (with an average of 20 blocks obtained across listeners and conditions) in mixed and pseudorandom order with threshold estimates computed from the last 10 blocks in each condition. Each block consisted of 60 trials, with a starting step size of 4 dB, reduced to 2 dB after the third reversal. The average of the reversals obtained was computed after discarding the initial three or four reversals depending on whichever produced an even number. The mean number of reversals obtained across all conditions was 10.2.

3. Results

The group mean results from all conditions are plotted in Fig. 1. First, all of the group-mean quiet signal thresholds—for both frequencies in single- and dual-signal conditions—fell within a very narrow range from 5.6 to 7 dB SPL. As may be seen in Fig. 1, the various maskers raised thresholds considerably for the single-tone conditions (compare open symbols in the left-hand panel with corresponding symbols in the right three panels). There was more masking (the difference in threshold between the quiet single-tone threshold and the masked single-tone threshold at each frequency) for the multitone conditions than for the corresponding noise masked conditions despite the fact that both were presented at roughly the same SPL, have similar “protected regions” that exclude masker components, and have the same distribution of expected level per masker component outside of the protected regions. As is apparent from the error bars in Fig. 1 the intersubject variability in the amount of masking is substantially greater for the informational masker than for the energetic masker, a finding that is common in the masking literature (cf. Kidd *et al.*, 2008). Note that the asynchronous multitone masker produced only a few decibels less masking than the synchronous masker. This will be discussed further in the following. Across masker conditions, the low- and high-frequency thresholds are very similar for the quiet and noise conditions, but are somewhat higher for the high-frequency tone than the low-frequency tone in the two multitone maskers.

A three-way within-subjects analysis of variance was conducted on the results with the main factors being masker condition (none, noise, asynchronous multitone, and synchronous

Fig. 1. Group mean thresholds (dB SPL), and intersubject standard errors of the means, for all conditions. The different panels are for the various quiet and masked conditions (quiet, noise, asynchronous and synchronous, left to right). Within each panel, thresholds for low- and high-frequency signals are plotted as circles and triangles, respectively. Open symbols represent results from single-signal conditions and closed symbols represent results from dual-signal conditions.

multitone), number of signals (single or dual), and signal frequency (low or high). Masker condition [$F(3,15)=50.8, p<0.001$] and number of signals [$F(1,5)=34.4, p<0.002$] were both significant, whereas signal frequency [$F(1,5)=0.2, p<0.66$] was not. The interaction of masker and number of signals was significant [$F(3,15)=4.1, p<0.03$] as was the number of signals with signal frequency [$F(3,15)=4.1, p<0.03$], whereas masker condition by signal frequency was not. The three-way interaction was also not significant.

The costs associated with the performance in dual-signal conditions may be estimated directly by taking the differences between single-signal and dual-signal thresholds. Those cost estimates are displayed in Fig. 2 as group mean differences and standard errors of the mean differences. For the quiet condition, the costs of dual-signal presentation were negligible, on average 0.4 and 1.1 dB for low and high signal frequencies, respectively. The dual-signal costs increased for the noise masker, with group mean differences of 5.1 dB for the low-frequency signal and 3.5 dB for the high-frequency signal. These values are quite consistent with those in the literature, especially considering that an identification judgment was required (i.e., high versus low signal frequency; cf. Green, 1961). For the multitone masker, the dual-signal costs were 11.2 and 4.4 dB for the low- and high-frequency tones, respectively, in the asynchronous condition and 13.6 and 6.6 dB in the synchronous condition. The apparently greater costs for the low-frequency signal in masked conditions have not, to our knowledge, been reported previously. There is no obvious reason for this trend, which deserves further study.

4. Discussion

The results of the current experiment demonstrate that substantial costs of monitoring multiple frequency channels may be observed under some conditions. These large effects—up to about

Fig. 2. Group mean differences in thresholds between single-signal and dual-signal tasks (dual-signal costs). The error bars are standard errors of the mean differences.

14 dB—are consistent with related work using uncertain frequency targets embedded in randomized multitone maskers reported by Richards and Neff (2004), although estimates of costs were not directly obtained in that study. In contrast to the relatively small elevation in threshold typically reported for detecting a tone in Gaussian noise when there is trial-by-trial uncertainty about the signal frequency (roughly 3–5 dB), the same signal frequency uncertainty often produced significantly greater elevations in threshold for a random-frequency multitone masker. And, somewhat unexpectedly, this same degree of signal frequency uncertainty produced almost no elevation in threshold when there was no masker present. Although the asynchronous multitone masker condition was included with the expectation that it would produce substantially less informational masking, this was not the case. The small effect of delaying the signal relative to the masker was somewhat surprising given our informal listening while deciding on signal parameters and past studies using onset asynchrony to reduce informational masking. Other studies (e.g., Neff, 1995; Durlach *et al.*, 2005; Leibold and Neff, 2007) reported release from informational masking due to onset asynchrony that was much larger than that found here. The smaller effect in the current study is probably due to the briefer delay used (10 ms), which apparently was not sufficient to perceptually segregate the signal from the masker very effectively for these listeners.

One alternative explanation for the current findings is that the dual-signal costs increase simply as the amount of masking increases without regard to the presence/absence or type of masker. This seems to be highly unlikely given the fairly large body of past work on uncertain signal frequency detection, which has typically used noise levels producing much more masking than our double-notched noise did yet the costs were about the same. For example, Green (1961) used a 40 dB spectrum level noise with threshold values around E/N_0 of 10–20 dB (signal duration was a variable), which, as noted in the text, compare favorably to our results albeit for significantly lesser amounts of masking. Furthermore, Creelman (1960) measured the effects of uncertain frequency signal detection for several levels of the background noise. His results, which were given in percent correct performance, also fail to show any systematic change in costs with noise level (and therefore amount of masking). Thus, this does not appear to be a concern in the context of the current study.

The current results pose significant problems for models that attempt to account for the detrimental effects of detecting a tone of uncertain, as compared to certain, frequency. First of all, as discussed in Sec. 1, a common characteristic of these models is that the variance in the decision statistic, which accounts for the elevation in threshold due to monitoring multiple channels, only takes into account the channels that could potentially contain the target. The findings of this study are incompatible with this fundamental assumption underlying such models. In all three cases tested here: no masker, double-notch-filtered noise, and multitone masker, the increase in the number of potential signal channels upon which costs were computed was the same; i.e., from one channel to two. However, the range of costs actually observed across these conditions was more than 13 dB and was closely tied to the specific condition tested. Very small costs (1 dB or less) were observed when no masker was present. This finding was unexpected based on the common assumption that quiet threshold is determined by internal noise that is considered to be largely independent across channels. If this were true the costs of increasing the number of channels should be very much the same as occurs for actual (externally applied) noise. Historically, and also in the noise condition tested here, the costs are about 3–5 dB. Furthermore, and even more difficult to explain, the very large costs (up to about 14 dB) found for the multitone masker means that there is an interaction between the number of signal channels that must be monitored and the type of masking (i.e., energetic versus informational) produced by the masker.

Informational masking, at least for the task of detecting tones in a multitone masker, usually is thought of as reflecting a failure of the listener to selectively attend to a single frequency channel and ignore or suppress the outputs of other irrelevant (masker) channels (cf. Kidd *et al.*, 2008). The finding here of much greater costs for the informational masker than for the energetic masker suggests that the ability to ignore irrelevant channels becomes much less effective when attention must be divided across two signal channels.² It is not clear how to

modify the existing models of signal frequency uncertainty to account for the variation in costs across masker conditions found in this study. Presumably, however, the decision rule must take into account events occurring in nonsignal channels as well as potential signal channels. The multitone masker obviously produces large fluctuations in the outputs of nonsignal frequency channels when considered across stimulus presentations. These fluctuations, or the varying perceptual salience of the individual masker tones or timbre of the complex as a whole, not only produce informational masking but may interact with signal frequency uncertainty to increase the costs of dividing attention across locations in frequency. Buus *et al.* (1986) summarized their findings for uncertain tone detection in noise by stating "... that listeners may not make an optimally weighted sum of independent stochastic variables from each channel, but may make only an unweighted sum across an optimally chosen set of channels." The current results may indicate that in some highly uncertain masking conditions listeners are prevented from choosing those channels optimally.

Another difficulty for models of uncertain signal frequency effects is that, in the design of this experiment, a "switching" strategy is available to the listener. In that case the listener could simply attend to one frequency channel in the first interval, make a single YES–NO detection decision, and then switch attention to the other frequency channel for the second interval and make a second YES–NO detection decision and then report the two decisions. This strategy should greatly reduce any two-channel monitoring costs. This is further evidence for less-than-ideal performance in these tasks. The listeners were fully informed about the structure of the experiment and the nature of the stimulus. Whether further training specifically intended to promote a channel-switching strategy would reduce the costs observed here is not currently known.

5. Summary

To summarize, when listeners were required to monitor and respond to events in two widely spaced frequency channels within each trial, detection thresholds were significantly higher than corresponding single-signal thresholds. However, the costs associated with dual-signal presentation were only observed in masked conditions and not in quiet listening. The magnitude of the effect depended upon the type of masker that was present, and consequently the type of masking that was produced, and was significantly higher when the masker was a random-frequency multitone complex than a notched-filtered noise.

Acknowledgments

This work was supported by AFOSR Award No. FA9550-05-1-2005 and by Grant No. DC004545 and DC02012 from NIH/NIDCD. F.G. was supported by F32 DC006526 from NIDCD.

References and links

¹In the present context, we are using the word "cost" to reflect an increase in detection threshold due to signal frequency uncertainty. The word cost sometimes is used in different ways such as in experiments in which costs and benefits are manipulated directly as controlled experimental variables.

²An alternative possibility is that the difference in costs observed for the noise versus multitone maskers reflects differences in the slopes of the underlying psychometric functions for a constant change in d' . Psychometric functions were not obtained in this study so the issue must await future work for clarification.

Buus, S., Schorer, E., Florentine, M., and Zwicker, E. (1986). "Decision rules in detection of simple and complex tones," *J. Acoust. Soc. Am.* **80**, 1646–1657.

Creelman, C. D. (1960). "Detection of signals of uncertain frequency," *J. Acoust. Soc. Am.* **32**, 805–810.

Durlach, N. I., Mason, C. R., Gallun, F. J., Shinn-Cunningham, B., Colburn, H. S., and Kidd, G., Jr. (2005). "Informational masking for simultaneous nonspeech stimuli: Psychometric functions for fixed and randomly mixed maskers," *J. Acoust. Soc. Am.* **118**, 2482–2497.

Green, D. M. (1961). "Detection of auditory sinusoids of uncertain frequency," *J. Acoust. Soc. Am.* **33**, 897–903.

Kidd, G., Jr., 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 Media, New York), pp. 143–190.

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

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

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

Neff, D. L., and Green, D. M. (1987). “Masking produced by spectral uncertainty with multicomponent maskers,” *Percept. Psychophys.* **41**, 409–415.

Richards, V. M., and Neff, D. L. (2004). “Cuing effects for informational masking,” *J. Acoust. Soc. Am.* **115**, 289–300.

Sorkin, R. D., Pohlmann, L. D., and Gilliom, J. D. (1973). “Simultaneous two channel signal detection. III. 630 and 1400 Hz signals,” *J. Acoust. Soc. Am.* **53**, 1045–1050.

Tanner, W. P., Swets, J. A., and Green, D. M. (1956). “Some general properties of the hearing mechanism,” Technical Rep. No. 30, Electronic Defense Group, University of Michigan, Ann Arbor, MI [cited in *Signal Detection Theory and Psychophysics*, D. M. Green and J. A. Swets (Kreiger, New York, 1974)].

Veniar, F. A. (1958). “Signal detection as a function of frequency ensemble. I,” *J. Acoust. Soc. Am.* **30**, 1020–1024.