LETTERS TO THE EDITOR

This Letters section is for publishing (a) brief acoustical research or applied acoustical reports, (b) comments on articles or letters previously published in this Journal, and (c) a reply by the article author to criticism by the Letter author in (b). Extensive reports should be submitted as articles, not in a letter series. Letters are peer-reviewed on the same basis as articles, but usually require less review time before acceptance. Letters cannot exceed four printed pages (approximately 3000–4000 words) including figures, tables, references, and a required abstract of about 100 words.

Frequency discrimination under conditions of comodulation masking release (L)

Emily Buss,^{a)} John H. Grose, and Joseph W. Hall III

Division Otolaryngology/Head and Neck Surgery, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599

(Received 30 March 2011; revised 27 January 2012; accepted 1 February 2012)

Masked detection thresholds can often be improved by introducing coherent masker amplitude modulation across frequency, a phenomenon referred to as comodulation masking release (CMR). While CMR can be large for detection, it is smaller for supra-threshold tasks, such as intensity discrimination. In this experiment, frequency discrimination for a 1000-Hz tone near threshold was found to be poorer in an amplitude-modulated than a steady bandpass noise. These results parallel previous findings for intensity discrimination. Although this study examined the relatively simple task of frequency discrimination, the results may have implications for more complex tasks, such as speech recognition in fluctuating noise. © 2012 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.3688508>]

PACS number(s): 43.66.Dc [LD]

Pages: 2557–2560

I. INTRODUCTION

Detection threshold for a tone masked by a random noise band can be reduced by the inclusion of comodulated flanking noise bands or by the introduction of coherent envelope modulation to a wideband masker (Hall *et al.*, 1984). This result has been described as comodulation masking release (CMR). There is a growing body of evidence that CMR is larger for detection than for supra-threshold discrimination. The reduction in CMR for supra-threshold tasks has been shown for intensity discrimination, gap detection, pitch ranking, melody recognition, and speech recognition (for reviews, see Buss and Hall, 2009; Hall *et al.*, 2011).

The reason for reduced CMR in supra-threshold tasks is unknown. One possibility proposed by Hall and Grose (1995) is that derived detection cues, such as those based on across-channel comparisons, may be of sufficient quality to support detection, but insufficient for fine discriminations. For example, if signal detection in a comodulated masker is based on the output of a modulation filterbank (Verhey *et al.*, 1999), then temporal fine-structure information for a signal near threshold may not be available to the listener. Recent data from Buss and Hall (2009) failed to find evidence of fundamentally different supra-threshold processing in steady and comodulated noise, however, casting doubt on the idea that derived cues in CMR reduce discrimination

abilities. In that study, pure-tone intensity discrimination for a tonal target in a noise masker was poorer for a target near threshold in a comodulated than a steady noise masker. Results of additional stimulus manipulations indicated that this pattern of results was likely due to the target-to-masker ratio rather than the presence or absence of across-channel cues. It was argued that fluctuation of the target-plus-masker disrupts intensity discrimination; for a tonal target in a noise masker, this effect is seen predominantly at low target-to-masker ratios due to the dominant effects of the noise on the envelope of the summed stimulus.

Whereas stimulus level fluctuation is known to interfere with intensity discrimination (Bos and de Boer, 1966), it can also reduce sensitivity to changes in frequency (Grant, 1987). Stimulus level affects perceived pitch for stationary stimuli (for a review, see Fastl and Zwicker, 2007), and randomizing presentation level across intervals has been shown to elevate frequency discrimination thresholds in some cases (Henning, 1966; Emmerich *et al.*, 1989). Partially masking a tone can also affect its perceived pitch (Burns and Oesterle, 1980; Fastl and Zwicker, 2007). These observations raise the possibility that frequency discrimination for a tonal target in a comodulated noise masker may be degraded due to stimulus envelope fluctuation, similar to the effects observed for intensity discrimination. It is also possible that the inherent frequency fluctuation of a noise masker could affect the pitch of the target, analogous to the effects observed for intensity discrimination.

Previous work has shown relatively poor supra-threshold pitch ranking (Hall *et al.*, 1997) and melody

^{a)}Author to whom correspondence should be addressed. Electronic mail: ebuss@med.unc.edu

discrimination (Hall *et al.*, 2011) for tones presented in comodulated noise. One possibility is that poor performance in comodulated noise for these tasks is due to poor frequency discrimination. However, both tasks are relatively complex, relying heavily on memory for multiple pitches, and no data on frequency discrimination are currently available. The goal of the present experiment was, therefore, to measure frequency discrimination thresholds in conditions associated with CMR. If stimulus fluctuation of the target-plus-masker disrupts the representation of pitch in the auditory system, then frequency discrimination for a pure tone in a fluctuating noise masker should be more adversely affected at low, rather than high, target-to-masker ratios.

II. EXPERIMENT

A. Methods

1. Observers

Observers were ten normal-hearing adults, ages 20–55 yrs (mean 33 yrs). All had thresholds of 15 dB hearing level (HL) or better for pure tones 250–8000 Hz (ANSI, 2004). None of these observers reported a history of ear disease, and all had previously participated in psychoacoustic experiments. Two additional observers meeting these inclusion criteria were recruited, but later excluded, based on poor performance and excessive variability in frequency discrimination performance. Both of these listeners had discrimination thresholds of greater than 5% in some conditions and provided sequential thresholds in a single condition that varied by a factor of four or more.

2. Stimuli

The target was a 1000-Hz pure tone. Targets were gated on and off with 50-ms raised-cosine ramps, and had a total duration of 400 ms. The masker was either a steady noise or a noise that had been amplitude modulated via multiplication with the Hilbert envelope of a 20-Hz wide, narrowband noise. These are referred to as the *steady* and *AM* masker conditions, respectively. In both cases the masker was a bandpass noise (190–1810 Hz) that played continuously at 65 dB sound pressure level (SPL). In addition to these masker conditions, thresholds were also collected for a target presented in quiet. The stimulus conditions of the present experiment closely resemble the “bandpass masker” conditions in experiment 1 of Buss and Hall (2009), which reported supra-threshold intensity discrimination. This similarity allows a comparison of supra-threshold intensity and frequency discrimination in *steady* and *AM* noise.

Noise samples associated with the bandpass masker and the narrowband noise modulator were generated in the frequency domain, with draws from a normal distribution defining the real and imaginary components within the associated passband. These arrays were transformed into the time domain, resulting in an array composed of 2^{17} points. When played out at 12 207 Hz, noise samples repeated seamlessly once every 10.7 s. New masker samples were generated in MATLAB prior to each threshold estimation run. All stimuli were played through a real-time digital processor (RP2,

TDT), passed through a headphone buffer (HB7, TDT), and presented to the left channel of a pair of circumaural headphones (Sennheiser, HD 265).

3. Procedures

Stimuli were presented using a three-alternative forced-choice procedure, with 300-ms inter-stimulus intervals. Thresholds were estimated using a 3-down 1-up tracking rule estimating 79% correct (Levitt, 1971). Either the level or the frequency of the target was adjusted in the course of a threshold estimation track. For the detection task, the target level was defined in units of dB SPL; initial target level adjustments were made in steps of 4 dB, and steps were reduced to 2 dB after the second track reversal. In the frequency discrimination task, the center frequency associated with the target was adjusted in factorial steps; at the outset of a track the target was adjusted by a factor of $2^{0.5}$, and this was reduced to $2^{0.25}$ after the second track reversal. For both detection and frequency discrimination tasks, a total of eight reversals was obtained in each track, two with large steps and six with small steps, and thresholds were based on the signal characteristics at the last six reversals. For detection, the threshold was the mean target level, and for frequency discrimination it was the geometric mean of the target frequency divided by the standard frequency of 1000 Hz.

Lights on a handheld response box indicated the three listening intervals and provided correct-answer feedback. In the detection task the target was presented in a randomly selected interval, and the observer indicated which interval contained the target by pressing the associated button on the response box. In the discrimination task there was a target presented in all three intervals. That target was centered on the standard (1000-Hz) frequency in two intervals, and in one interval its center frequency was higher than 1000 Hz. The observer selected the interval with the higher-frequency target. In frequency discrimination tracks, the level of the target was set relative to each observer’s masked detection threshold in the associated condition: standard levels were either 10, 20, or 30 dB sensation level (SL). For frequency discrimination in quiet, the target was presented at the level associated with 10- or 30-dB SL relative to detection threshold in the *steady* masker condition.

Detection conditions were completed first, with the order of masker conditions randomly selected for each observer. Three estimates were obtained in each condition, and a fourth estimate was obtained in cases where the initial estimates spanned a range of 3 dB or more. The resulting mean thresholds for each observer were then used to determine the signal presentation levels for frequency discrimination testing. Discrimination thresholds were obtained blocked by condition, with conditions run in quasi-random order.¹ Each block included at least three estimates, and in most cases a fourth estimate was obtained. A group of five listeners repeated the frequency discrimination conditions after completion of the experiment to assess possible effects of practice. These listeners completed the second set of discrimination thresholds in a new random order. The first set of data obtained on all listeners is reported below, and a

separate analysis evaluates the possible effect of practice for this subset of five listeners.

B. Results

Mean target detection thresholds were 51.9 dB SPL in the *steady* masker and 45.9 dB SPL *AM* masker, resulting in a CMR of 6.0 dB. A paired t-test confirmed that this masking release was significantly greater than zero ($t_9 = 21.163$, $p < 0.001$). These detection thresholds closely match those reported by [Buss and Hall (2009); 52.1 and 45.8 dB, respectively].

Frequency discrimination in quiet was relatively insensitive to presentation level. Mean frequency discrimination thresholds in quiet were 0.47% and 0.39% for the target levels of 10 and 30 dB SL relative to thresholds in the *steady* masker. This trend is consistent with the expectation of better performance at higher presentation levels. A repeated-measures analysis of variance (ANOVA) was performed on the log transform of individual observers' frequency discrimination thresholds in quiet, with two levels of stimulus LEVEL (10, 30 dB SL re *steady*). The effect of presentation LEVEL did not reach significance ($F_{1,9} = 1.982$, $p = 0.193$), so thresholds in quiet at the two levels were combined for plotting purposes.

Figure 1 shows geometric means of the masked frequency discrimination thresholds, plotted as a function of the signal level relative to detection threshold. Results for discrimination in quiet are shown at the far right of the panel, for reference. Error bars show ± 1 standard error the mean, computed in log units, and listening conditions are indicated with symbols, as defined in the legend. This figure shows that masked discrimination thresholds tended to improve with increasing target level in both masker conditions, approaching thresholds in quiet by 30 dB SL. Threshold reduction with increasing target level was more pronounced for the *AM* than the *steady* masker condition, however. This result resembles

FIG. 1. The geometric mean of frequency discrimination thresholds across observers, in units of percent of the standard (1000-Hz) frequency, plotted as a function of target level relative to detection threshold. Error bars indicate ± 1 standard error of the mean across the ten observers, computed in log units. Symbols reflect stimulus conditions, as indicated in the legend. The geometric mean of frequency discrimination thresholds for a tone in quiet is shown at the far right of the figure.

that obtained for intensity discrimination under comparable conditions, where the masker effect was largest for the 10-dB-SL target, and thresholds converged at the highest target level (Buss and Hall, 2009; Fig. 1, panel B).

The trend for larger level effects in the *AM* than the *steady* masker condition was confirmed by performing a repeated-measures ANOVA on the log-transformed thresholds, with two levels of COND (*AM*, *steady*) and three of LEVEL (10, 20, 30 dB SL). There was a main effect of LEVEL ($F_{2,18} = 30.950$, $p < 0.0001$) but not COND ($F_{1,9} = 4.076$, $p = 0.074$). The interaction between COND and LEVEL was significant ($F_{2,18} = 4.568$, $p = 0.025$). Pre-planned contrasts indicated that thresholds in the two masker conditions differed at 10 dB SL ($p = 0.013$), but not at 20 dB SL ($p = 0.095$) or 30 dB SL ($p = 0.344$).

Data for the five listeners who completed the experiment twice were evaluated for effects of practice, with a particular interest in possible effects of practice on the COND-by-LEVEL interaction. A repeated-measures ANOVA was performed on the log-transformed thresholds, with two levels of COND (*steady*, *AM*), three of LEVEL (10, 20, 30 dB SL), and two levels of ESTIMATE (first, second). This resulted in a significant effect of ESTIMATE ($F_{1,4} = 9.834$, $p = 0.035$), reflecting an improvement of approximately 0.2% in masked frequency discrimination between the first and second dataset. None of the interactions with ESTIMATE were significant, however. Most noteworthy, the three-way interaction failed to approach significance ($F_{1,4} = 2.19 \times 10^{-5}$, $p = 0.986$). This result indicates that whereas additional practice may have reduced thresholds, it is very unlikely to have changed the differential effect of level in the two masker conditions.

While the present data on frequency discrimination follow the same general trends as seen in the published intensity discrimination data, the masker effect appears to be less robust than that observed for intensity discrimination under comparable stimulus conditions. Frequency discrimination thresholds in percent cannot be directly compared to intensity discrimination thresholds in $10\log(\Delta I/I)$, but the magnitude of the masker effect can be compared to the magnitude of the variability across observers' threshold estimates. The magnitude of the masker-by-level interaction in the frequency discrimination data of Fig. 1 is modest relative to the associated error bars. In contrast, this interaction in the comparable figure of Buss and Hall (2009); [Fig. 1, panel B] for intensity discrimination is larger relative to the associated error bars.² This visual impression can be quantified by estimating effect size of the COND-by-LEVEL interaction, which was nearly a factor of 2 smaller in the present frequency discrimination data (partial $\eta^2 = 0.34$) than in the published intensity discrimination data [partial $\eta^2 = 0.62$; Buss and Hall (2009)].

C. Discussion

The present experiment was carried out to document CMR for frequency discrimination. Supra-threshold frequency discrimination for a pure-tone target presented near threshold was found to be poorer under conditions of

masking release than in baseline masking conditions at comparable levels relative to detection threshold (dB SL). This result is consistent with previous work showing relatively poor supra-threshold pitch ranking (Hall *et al.*, 1997) and melody discrimination (Hall *et al.*, 2011) for tones presented near threshold in comodulated noise. It is also consistent with data on supra-threshold intensity discrimination (Buss and Hall, 2009).

Buss and Hall (2009) argued that the masker effect for supra-threshold intensity discrimination with a tonal target in a comodulated masker was due in part to the level fluctuation of the target-plus-masker at low target-to-masker ratios. This could be related to the finding that inherent envelope modulation interferes with intensity discrimination for a narrowband noise stimulus (Bos and de Boer, 1966). It was suggested that frequency discrimination thresholds might also be elevated due to stimulus level fluctuation for a tonal target near threshold in a comodulated noise masker. Perceived pitch is affected by stimulus level, such that fluctuating level could introduce variability in perceived pitch. Studies of pitch perception in quiet indicate that level effects are frequency-specific, with more intense tones being associated with lower pitch at low frequencies and higher pitch at high frequencies (Morgan *et al.*, 1951). Some studies have shown more modest effects of level for tones at 1000 and 2000 Hz than for tones at lower or higher frequencies (Morgan *et al.*, 1951; Henning, 1966). For example, randomizing level has a much more pronounced effect on pitch for targets above than below 4000 Hz (Henning, 1966). It is therefore possible that the modest effects observed here for a 1000-Hz target would be larger for a lower or higher target frequency. However, mid frequencies are not always associated with smaller level effects (Emmerich *et al.*, 1989), and results at more extreme target frequencies would arguably be less relevant to speech perception in comodulated noise. Further, substantial individual differences in the effect of level on pitch exist even for extreme target frequencies, where some listeners show a large effect and others show little or no level effect (Morgan *et al.*, 1951).

The masker effect reported here for frequency discrimination appears to be more modest than that previously observed for intensity discrimination when compared against estimate variability. This could be due to modest effects of stimulus fluctuation on pitch at 1000 Hz, but it could also be related to the finding of larger individual differences for frequency than intensity discrimination (Jesteadt and Bilger, 1974) or to the large individual differences in the effect of level on pitch (Morgan *et al.*, 1951; Emmerich *et al.*, 1989; Dai *et al.*, 1995). While the effect of masker fluctuation is small for the stimuli used here, it is possible that this effect could play a substantial role in speech perception in modulated noise. Recent data on melody discrimination indicate that pitch information available to the listener for a tonal target in comodulated noise may be particularly poor for brief targets that are presented in quick succession (Hall *et al.*, 2011). The poor supra-threshold frequency discrimination observed here for a relatively long duration target could likewise be more pronounced for brief or dynamic stimulus features, such as those required for accurate speech perception.

The present study demonstrates that frequency discrimination is reduced for targets near threshold under conditions of CMR. In combination with previously reported supra-threshold deficits in intensity discrimination and gap detection, such supra-threshold effects could impact speech perception in comodulated noise. These effects could be particularly important to consider in hearing-impaired listeners, for whom temporal and/or spectral resolution may be impaired. In these listeners, the introduction of effects related to masker fluctuation may obfuscate the minimal cues required for speech recognition.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health, NIDCD: R01 DC000418 (JWH) and R01 DC007391 (EB).

¹Data were also collected with a narrowband noise target. Those data were inconclusive, however, and so have been omitted.

²Both the present report and that of Buss and Hall (2009) plotted standard error of the mean. Since there were fewer observers in the previous report ($n = 7$ vs $n = 10$), this comparison slightly underestimates the difference in effect size across experiments.

ANSI (2004). *ANSI S3.6-2004, American National Standard Specification for Audiometers* (American National Standards Institute, New York).

Bos, C. E., and de Boer, E. (1966). "Masking and discrimination," *J. Acoust. Soc. Am.* **39**, 708–715.

Burns, E. M., and Oesterle, E. L. (1980). "Pitch shifts associated with partial masking," *J. Acoust. Soc. Am.* **67**, S20–S20.

Buss, E., and Hall, J. W. I. (2009). "Effects of masker envelope coherence on intensity discrimination," *J. Acoust. Soc. Am.* **126**, 2467–2478.

Dai, H., Nguyen, Q. T., and Green, D. M. (1995). "A two-filter model for frequency discrimination," *Hear. Res.* **85**, 109–114.

Emmerich, D. S., Ellermeier, W., and Butensky, B. (1989). "A reexamination of the frequency discrimination of random-amplitude tones, and a test of Henning's modified energy-detector model," *J. Acoust. Soc. Am.* **85**, 1653–1659.

Fastl, H., and Zwicker, E. (2007). "Pitch and pitch strength," in *Psychoacoustics Facts and Models* (Springer, Berlin), pp. 111–148.

Grant, K. W. (1987). "Frequency modulation detection by normally hearing and profoundly hearing-impaired listeners," *J. Speech Hear. Res.* **30**, 558–563.

Hall, J. W., Buss, E., and Grose, J. H. (2011). "Masked detection and discrimination of tone sequences under conditions of monaural and binaural masking release," *J. Acoust. Soc. Am.* **129**, 1482–1489.

Hall, J. W., and Grose, J. H. (1995). "Amplitude discrimination in masking release paradigms," *J. Acoust. Soc. Am.* **98**, 847–852.

Hall, J. W., Grose, J. H., and Dev, M. B. (1997). "Signal detection and pitch ranking in conditions of masking release," *J. Acoust. Soc. Am.* **102**, 1746–1754.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (1984). "Detection in noise by spectro-temporal pattern analysis," *J. Acoust. Soc. Am.* **76**, 50–56.

Henning, G. B. (1966). "Frequency discrimination of random-amplitude tones," *J. Acoust. Soc. Am.* **39**, 336–339.

Jesteadt, W., and Bilger, R. C. (1974). "Intensity and frequency discrimination in one- and two-interval paradigms," *J. Acoust. Soc. Am.* **55**, 1266–1276.

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

Morgan, C. T., Garner, W. R., and Galambos, R. (1951). "Pitch and intensity," *J. Acoust. Soc. Am.* **23**, 658–663.

Verhey, J. L., Dau, T., and Kollmeier, B. (1999). "Within-channel cues in comodulation masking release (CMR): Experiments and model predictions using a modulation-filterbank model," *J. Acoust. Soc. Am.* **106**, 2733–2745.