Effects of flanking component spectral position and modulation pattern on thresholds for signals presented in the peaks of a modulated tonal masker

Roel Delahaye
Parmly Hearing Institute, Loyola University, Chicago, Illinois 60626

Deborah A. Fantini and Ray Meddis
Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, England

(Received 23 May 2000; revised 21 September 2000; accepted 28 September 2000)

Detection of signals added solely to the peaks of an on-frequency modulated masker has never been found to improve after adding flanking components (FCs) at remote frequencies. However, according to theories underlying comodulation masking release (CMR), adding comodulated FCs could provide cues that improve signal detection for signal-peak placement. A reason masking release was not found with peak placement might be because of processes underlying modulation detection interference (MDI). This possibility was further investigated by using FCs that could diminish MDI but still provide signal detection cues associated with theories of CMR. It seemed that a peripheral within-channel and a central across-channel mechanism underlying MDI could hinder signal detection for signal peak placement.
[DOI: 10.1121/1.1328790]

PACS numbers: 43.66.Dc, 43.66.Mk, 43.66.Rq [SPB]

I. INTRODUCTION

This study focuses on the ability of the auditory system to use signal detection cues predicted by theories of comodulation masking release (CMR) (Hall *et al.*, 1984). CMR is referred to as the detection improvement of a signal masked by a fluctuating masker (on-frequency component, OFC) after adding modulated components with the same envelope-phase relationship as the OFC to frequencies remote from the signal frequency (flanking components, FCs). One of the theories of CMR involves comparisons between masker envelopes. One aspect of the envelope comparison theory is based upon the difference in envelope amplitude, derived by subtracting the envelope at the signal frequency from the envelope at the flanking frequency (Buus, 1985; Hall, 1986).

Another theory of CMR is not based on a direct comparison between masker envelopes, but on listening for the signal during the OFC envelope when the signal-to-masker ratio is greatest. The signal-to-masker ratio is usually greatest around the OFC minima (dips). Thus, the more the dips in the OFC envelope are cued by FC dips, the more masking release there should be. This is called the dip-listening theory (Buus, 1985). An experiment conducted by Hall and Grose (1988) showed that CMR can be obtained for stimuli that contain cues based only on amplitude level differences across frequencies or contain only cues for dip listening.

The possibility of obtaining CMR for different signal detection cues is probably the reason that CMR can be found for a variety of stimuli (see also Schooneveldt and Moore, 1987; Fantini *et al.*, 1993; Fantini and Moore, 1994). Despite this, CMR has never been observed for signals placed solely in the peaks of a modulated OFC (Grose and Hall, 1989; Moore *et al.*, 1990). However, if the mechanism underlying CMR relied on an across-channel difference in envelope amplitude, some masking release would be expected even when the signal occurred only in the peaks of the masker. Moore *et al.* (1990) suggested that the reason why no masking release has been found for the comodulated signal peak-placement condition might be because of perceptual fusion of the FCs and the OFC plus signal. This fusion might be related to a mechanism underlying modulation detection interference (MDI) as described by Yost and Sheft (1989). Yost and Sheft found that if a listener had to detect the modulation of a target component in the presence of another component, it was harder to detect the modulation when the other component was also modulated.

The experiments in this study were designed to investigate if mechanisms similar to the ones underlying MDI might prevent the benefit to signal peak-placement detection of dip-listening and envelope-level cues associated with CMR. The issue of dip-listening cues was addressed by using FCs that were modulated 180° out of phase with the OFC (antiphasic FCs). The dips of the antiphasic FCs might indicate the optimum time to listen for the signal (the OFC peaks). This might give an advantage in signal detection over the no-FC condition in which no indication of the signal timing by FCs exists. The issue of envelope-level cues was addressed by using unmodulated FCs. No interference in signal detection due to common modulation of any of the masker components can take place. Therefore, across-frequency level disparities can be used as a detection cue without the mechanism underlying MDI impairing signal detection. To determine the role of peripheral and central processes, the FCs were presented ipsilaterally or contralaterally to the OFC and signal. In the first experiment, signal detection was measured as a function of the modulation pattern of six FCs. In the second experiment only one FC was used. Signal detection was measured as a function of the FC spectral position.

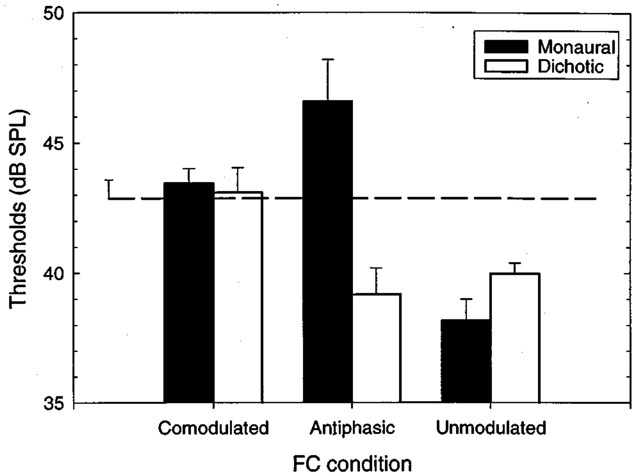

FIG. 1. The average results for monaural (filled columns) and dichotic (open columns) stimulus presentations are shown. The dashed line represents the average threshold obtained when no FCs were present. Signal detection thresholds are plotted for the different FC conditions. Error bars indicate plus one standard error of the mean.

II. METHOD

A. Listeners

Three normal-hearing listeners participated in this experiment. The first two listeners were the first two authors. The third listener was paid for her participation. All listeners were between the ages of 20 and 40 years. Data collection for experiment 1a started after a listener had at least 10 h of practice and showed a stable performance. Data for experiment 1b were collected after the listeners gained an additional amount (about 20 h) of practice by participating in similar psychoacoustical experiments. Listeners always ran individually in an IAC double-walled sound-attenuating chamber.

B. Stimuli

The stimuli were essentially the same as the ones used by Grose and Hall (1989) and Moore *et al.* (1990). The signal was a 700-Hz pure tone, which was presented as three 50-ms tone bursts, each with 20-ms raised-cosine rise and fall ramps and a steady-state duration of 10 ms. The time interval between the bursts was 50 ms. The on-frequency component (OFC) was a 700-Hz sinusoid which was 100% sinusoidally amplitude modulated (SAM) at a 10-Hz rate. The OFC had an overall duration of 500 ms, with 20-ms raised-cosine rise and fall ramps, and an overall level of 51.8 dB SPL after modulation. The signal was temporally centered on the middle three peaks of the OFC envelope and was presented in phase with respect to the OFC carrier.

Pure-tone flanking components (FCs) could be presented ipsilaterally or contralaterally to the OFC and signal. In experiment 1a six FCs were used. These FCs had carrier frequencies of 300, 400, 500, 900, 1000, and 1100 Hz. In experiment 1b a single FC was used. This FC had a carrier frequency of either 100, 300, 500, 600, 650, 750, 800, 900, 1100, or 1300 Hz. FC duration was the same as for the OFC. FCs were either modulated or unmodulated. Each unmodulated FC had an overall level of 50 dB SPL. Each modulated FC had the same level (51.8 dB SPL), SAM depth (100%), and modulation frequency (10 Hz) as the OFC. The modulation phase of the FCs with respect to the OFC envelope was varied in two different ways. First, in the comodulated condition, the FCs were modulated in phase with respect to the OFC. Second, in the antiphasic condition, the FC envelopes were 180° out of phase with the OFC envelope. A condition in which no FCs but only the OFC was present was also run. This no-FC condition was run in both experiments to account for differences in performance due to an unequal amount of practice preceding the data collection in experiments 1a and 1b.

A NeXT work station generated the stimuli digitally using a 44.1-kHz sampling rate, converted them using a 16-bit D/A, and filtered them using an antialias filter. Stimuli were presented through Sennheiser HD 340 headphones.

C. Procedure

A three-interval forced-choice adaptive-tracking paradigm, with a two-down, one-up strategy (Levitt, 1971) was used to determine the 71% correct signal detection thresholds in dB SPL. The starting level of the signal was always above threshold and the step size was 5 dB for the first four reversals. For the next eight reversals, a step size of 2 dB was used. Data collected for the first four reversals were omitted and signal levels for the last eight reversals were averaged to estimate threshold. Runs in which the standard deviations across eight reversals exceeded 5 dB were discarded and repeated. Estimates of three signal thresholds were collected for each condition. If the range of the three thresholds exceeded 3 dB another estimate was obtained. The final threshold value was the average of all the threshold estimates for each condition. A warning light preceded each block of trials. Each observation interval was indicated by a light, followed by a 500-ms silent interval. After a response was given and visual feedback was provided the next trial started immediately.

III. RESULTS

A. Experiment 1a

Signal thresholds of the three listeners were averaged and plotted in Fig. 1 for each FC condition. Filled columns represent thresholds obtained in the monaural condition and open columns represent thresholds obtained in the dichotic conditions. The dashed line represents the threshold obtained for the no-FC condition. Bars indicate plus one standard error of the mean.

Thresholds obtained in the comodulated, antiphasic, and unmodulated conditions were considered with respect to the condition in which no FCs were present. Adding comodulated FCs did not seem to have an effect on signal detection. The presence of antiphasic FCs impaired signal detection in the monaural condition but improved signal detection in the dichotic condition. Adding unmodulated FCs resulted in a signal detection improvement for both monaural and dichotic stimulus presentations.

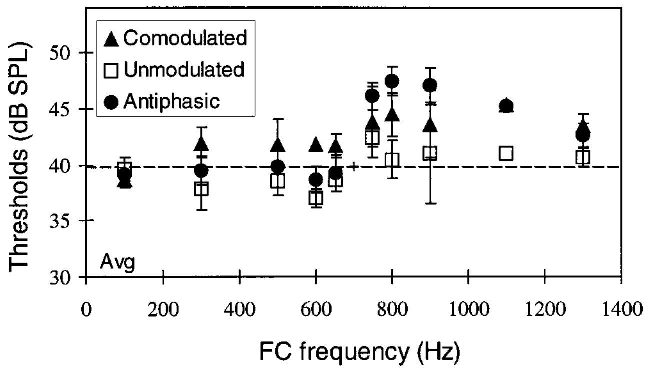

FIG. 2. The monaural average results are shown. Signal detection thresholds are plotted as a function of FC frequency. Triangles represent the comodulated conditions, circles represent the antiphasic conditions, and squares represent the unmodulated conditions. The horizontal dashed line represents the signal threshold obtained in the no-FC condition. Bars represent plus and minus one standard error of the mean. Error bars are, for clarity's sake, omitted when they did not exceed the symbol size.

B. Experiment 1b

The average results are shown in Fig. 2 for the monaural conditions. Figure 3 shows the individual results for the dichotic condition. Signal threshold is plotted as a function of the FC frequency. Triangles represent the comodulated conditions, circles represent the antiphasic conditions, and squares represent the unmodulated conditions. The horizontal dashed line represents the signal threshold obtained for the no-FC condition. Bars represent plus and minus one standard error of the mean. Error bars are, for clarity's sake, omitted when they did not exceed the symbol size.

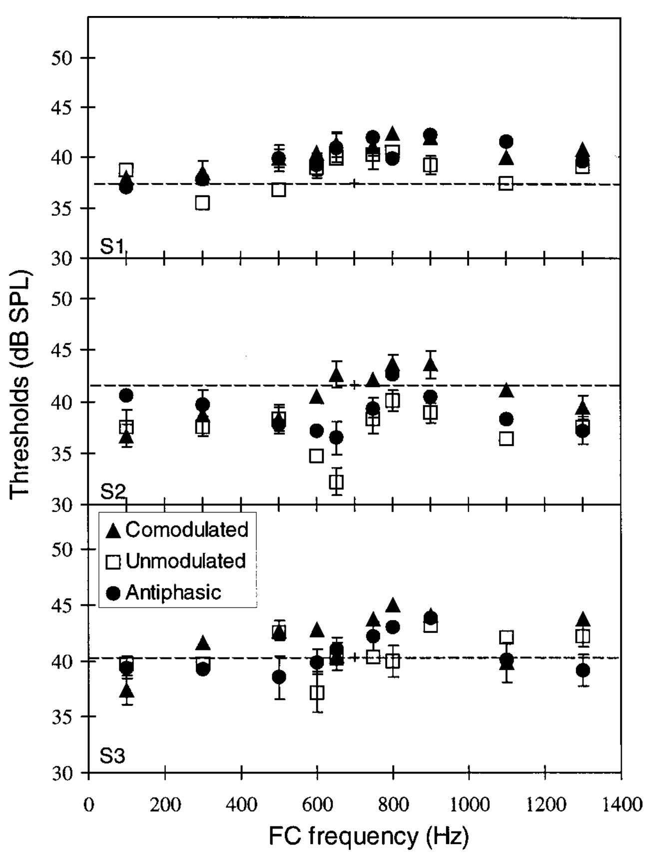

FIG. 3. As Fig. 2, but for dichotic results for the individual listeners.

First, consider thresholds obtained for the monaural conditions (Fig. 2). Only the average results are shown as the threshold patterns among individual listeners were similar. For four out of five FCs positioned below the signal frequency, the lowest average thresholds were found for the unmodulated conditions followed by the antiphasic and comodulated conditions, respectively. The threshold obtained for the no-FC condition was similar to the thresholds obtained for the antiphasic conditions. When FCs were placed above the signal frequency, thresholds were higher than those found for the no-FC condition. The lowest thresholds were found for the unmodulated conditions followed by the comodulated and antiphasic conditions, respectively. For dichotic conditions (Fig. 3), only the individual results are shown as they vary substantially among listeners. A clear distinction between the thresholds obtained for the different conditions could not be made.

For both the unmodulated and comodulated conditions, there was no clear difference between the monaural and dichotic conditions. However, for the antiphasic conditions, the difference was rather striking. Antiphasic monaural and dichotic thresholds were similar when FCs were placed below the signal frequency, but when antiphasic FCs were placed above the signal frequency, dichotic thresholds were consistently lower than monaural thresholds.

IV. DISCUSSION

Thresholds obtained in the no-FC condition of experiment 1b were on average about 3 dB lower than in experiment 1a. This might be due to additional practice obtained before commencing the second experiment. Although performance appeared stable after the initial 10 h of practice, it seemed that another 20 h of experience could still decrease thresholds in the no-FC condition. Delahaye *et al.* (1999) measured the effect of extended practice in the six-FC conditions for two of the three listeners used here. They showed that, after extended practice, thresholds obtained in the no-FC condition seem to decrease more than thresholds obtained in the FC conditions. Extended practice reduced or diminished masking release when measured with respect to the no-FC condition. However, the threshold difference between FC conditions seemed independent of practice. Therefore, Delahaye *et al.* (1999) proposed using a reference condition that also contains FCs. This proposal is taken into account in this study by not only discussing FC conditions in respect to the no-FC condition but also in respect to each other.

The lack of signal detection improvement after adding comodulated FCs (CMR) is in agreement with the findings of Moore *et al.* (1990) and Grose and Hall (1989). In fact, they showed a tendency for an increase in thresholds after adding comodulated FCs to the OFC. In experiment 1a, most similar to the experiments by Moore *et al.* (1990) and Grose and Hall (1989), such an increase was not found. However, a closer look at the results obtained in the previous studies showed that, although higher thresholds were generally found for the comodulated than for the no-FC conditions, this difference was not significant in the study by Grose and Hall (1989) and not significant in the study by Moore *et al.*

(1990) for the conditions most similar to the ones used in experiment 1a (seven-component masker). Thus, the current results are in general agreement with the previous results. If the lack of a masking release in the comodulated condition is due to a mechanism that hinders across-frequency level comparisons, this mechanism is probably involved with central across-channel interactions as no masking release is found in both monaural and dichotic conditions.

In the antiphasic condition of experiment 1a, the release from masking found in the dichotic condition is in contrast with the increase in masking found in the monaural condition. The masking release found in the dichotic condition indicates that it is produced by an across-channel mechanism. This mechanism might be related to the dip-listening theory (Buus, 1985). Antiphasic FC dips might indicate the optimal time to listen for the signal during the OFC peaks where the signal is present. This might give an advantage in signal detection over the no-FC condition. For the monaural condition it seems that the effect of dip listening is counteracted by peripheral within-channel interactions. From experiment 1b it can be seen that these peripheral within-channel interactions occur mainly for FCs placed immediately above the signal frequency.

The lack of a masking release for the monaural and dichotic comodulated conditions and the monaural antiphasic condition might be related to the MDI phenomenon as discussed by Moore and Jorasz (1996). In a monaural setting they showed that MDI could be due to both peripheral within-channel and central across-channel interactions. They argued that the MDI caused by peripheral within-channel interactions was greater when the target and flanker were antiphasic than when they were comodulated. Conversely, the MDI caused by central across-channel interactions was supposed to be greater in the comodulated than in the antiphasic condition. Also, Moore and Jorasz (1996) argued that for FCs placed above the target frequency, peripheral interactions were dominant, resulting in more MDI for the antiphasic than for the comodulated condition. For FCs placed below the signal frequency, central processes were dominant resulting in more MDI for the comodulated than for the antiphasic condition. Notice that Moore and Jorasz (1996) used only monaural stimulus conditions. However, in this study both monaural and dichotic conditions were used. It was assumed that peripheral interactions could occur only in the monaural condition as central interactions could occur in both the monaural and dichotic conditions. Therefore, the peripheral mechanism underlying MDI in the antiphasic condition is assumed to operate only in the monaural condition. In contrast, the central mechanism underlying MDI in the comodulated condition is assumed to operate in both the monaural and dichotic conditions. Therefore, the MDI effect caused by the central mechanism could be of the same magnitude in the monaural and dichotic conditions. The similarity in threshold pattern between the current experiments and the MDI experiment by Moore and Jorasz indicates that mechanisms underlying MDI might hinder signal detection for signal peak placement.

The lowest thresholds were generally found for the unmodulated conditions. In experiment 1a, adding unmodulated FCs resulted in a masking release for both the monaural and dichotic conditions. It seemed that in the unmodulated condition, MDI could not hinder signal detection as it was argued to do in the comodulated and monaural antiphasic conditions. Therefore, in the unmodulated condition across-frequency level comparisons could be used as a detection cue without the mechanism responsible for MDI impairing signal detection.

V. CONCLUSIONS

A reason why CMR has never been obtained for signal-peak placement conditions might have been due to an interfering, central across-channel mechanism associated with MDI. However, a theory associated with CMR based on amplitude level comparisons across frequencies still applies for peak-placement conditions when MDI is diminished by using unmodulated FCs. Support for another theory associated with CMR, the dip-listening theory, is found for results obtained in the dichotic antiphasic peak-placement condition. A reason no support for the dip-listening theory is found in the monaural antiphasic peak-placement condition might be due to a peripheral within-channel process underlying MDI.

Buus, S. (**1985**). ''Release from masking caused by envelope fluctuations,'' J. Acoust. Soc. Am. **78**, 1958–1965.

Delahaye, R., Fantini, D. A., and Meddis, R. (**1999**). ''Effect of practice on performance for different masking tasks,'' J. Acoust. Soc. Am. **105**, 1153.

Fantini, D. A., and Moore, B. C. J. (**1994**). ''A comparison of the effectiveness of across-channel cues available in comodulation masking release and profile analysis tasks,'' J. Acoust. Soc. Am. **96**, 3451–3462.

Fantini, D. A., Moore, B. C. J., and Schooneveldt, G. P. (**1993**). ''Comodulation masking release as a function of type of signal, gated or continuous masking, monaural or dichotic presentation of flanking bands, and center frequency,'' J. Acoust. Soc. Am. **93**, 2106–2115.

Grose, J. H., and Hall, J. W. (**1989**). ''Comodulation masking release using SAM tonal complex maskers: Effects of modulation depth and signal position,'' J. Acoust. Soc. Am. **85**, 1276–1284.

Hall, J. W. (**1986**). ''The effect of across-frequency differences on masking level on spectro-temporal pattern analysis,'' J. Acoust. Soc. Am. **79**, 781–787.

Hall, J. W., and Grose, J. H. (**1988**). ''Comodulation masking release: Evidence for multiple cues,'' J. Acoust. Soc. Am. **84**, 1669–1675.

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.

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

Moore, B. C. J., Glasberg, B. R., and Schooneveldt, G. P. (**1990**). ''Across-channel masking and comodulation masking release,'' J. Acoust. Soc. Am. **87**, 1683–1694.

Moore, B. C. J., and Jorasz, U. (**1996**). ''Modulation discrimination interference and comodulation masking release as a function of the number and spectral placement of narrow-band noise modulators,'' J. Acoust. Soc. Am. **100**, 2373–2381.

Schooneveldt, G. P., and Moore, B. C. J. (**1987**). ''Comodulation masking release (CMR): Effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band,'' J. Acoust. Soc. Am. **82**, 1944–1956.

Yost, W. A., and Sheft, S. (**1989**). ''Across-critical-band processing of amplitude-modulated tones,'' J. Acoust. Soc. Am. **85**, 848–857.