Detection of modulation of a 4-kHz carrier (L)

Neal F. Viemeister, Mark A. Stellmack, and Andrew J. Byrne

Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455

(Received 14 May 2009; revised 12 July 2010; accepted 13 July 2010)

To better understand the processing of complex high-frequency sounds, modulation-detection thresholds were measured for sinusoidal frequency modulation (SFM), quasi-frequency modulation (QFM), sinusoidal amplitude modulation (SAM), and random-phase FM (RPFM). At the lowest modulation frequency (5 Hz) modulation thresholds expressed as AM depth were similar for RPFM, SAM and QFM suggesting the predominance of envelope cues. At the higher modulation frequencies (20 and 40 Hz) thresholds expressed as total frequency excursions were similar for SFM and QFM suggesting a common mechanism, one perhaps based on single-channel FM-to-AM conversion or on a multi-channel place mechanism. The fact that the nominal envelopes of SFM and QFM are different (SFM has a flat envelope), seems to preclude processing based on the envelope of the external stimulus. Also, given the 4-kHz carrier and the similarity to previously published results obtained with a 1-kHz carrier, processing based on temporally-coded fine structure for all four types of modulation appears unlikely.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3474897]

PACS number(s): 43.66.Nm, 43.66.Mk [BCM]

Pages: 1578–1581

I. INTRODUCTION

A useful way to characterize sounds in aspects that are relevant to auditory perception is in terms of the envelope and instantaneous frequency. This is especially true for sounds that have a narrow bandwidth relative to the filtering that occurs in the auditory system. There has been extensive research aimed at distinguishing the relative importance of such cues in auditory perception and more generally how such cues are processed. Typically this research has used canonical stimuli such as sinusoidal amplitude modulation, sinusoidal frequency modulation and relatively simple stimuli such as quasi-frequency modulation that contain both envelope and instantaneous frequency fluctuations (see [Edwards and Viemeister, 1994](#); [Moore and Sek, 1996](#)). Here we extend this research by using a high carrier frequency (4 kHz), introducing a stimulus that contains complex envelope and frequency fluctuations, and analyzing the data in terms of the envelope and frequency changes present in the stimulus. In the present experiment modulation detection thresholds for the following signals were measured:

- Sinusoidal FM (SFM). This signal has a flat envelope and periodic (sinusoidal) changes in instantaneous frequency.
- Random-phase FM (RPFM). This signal is derived by randomizing the starting phases of the components of its parent SFM signal. Its envelope and instantaneous frequency changes are periodic but are random across presentations.
- Sinusoidal AM (SAM). This signal has a periodic (sinusoidal) envelope and no changes in instantaneous frequency.
- Quasi FM (QFM). This is derived from its parent SAM by rotating the relative phase of the component at the carrier frequency by 90 degrees. Its envelope is periodic with a fundamental frequency that is twice that of the modulation frequency of the parent SAM but with reduced modulation depth. Its instantaneous frequency changes periodically.

It is important to note that these characterizations of the signals pertain to the stimulus itself, its “external” properties. After cochlear processing these properties of the signals may be considerably altered. Indeed, an account for the detection of FM is “FM-to-AM conversion” (e.g., [Zwicker, 1962](#)): As the instantaneous frequency sweeps through the passband of an auditory filter the output of the filter will change and thus frequency changes can be represented as amplitude changes within a single auditory filter.

Part of the rationale for investigating RPFM is that in our studies of FM processing RPFM was used in a control condition to evaluate the possibility that modulation detection of SFM was based on detection of the sidebands produced by modulation. RPFM is constructed so that its amplitude spectrum is identical to that for SFM. Modulation thresholds were considerably lower for RPFM despite its stochastic nature. The question is why human performance is considerably better for RPFM than for SFM. The likely hypothesis is that the envelope fluctuations in the external stimulus, although random, mediate detection of RPFM and provide a better cue than the regular instantaneous frequency changes in SFM.

The general purpose of this experiment is to distinguish between the use of envelope and instantaneous frequency cues in the detection of modulation produced by the four signals described above. The analysis technique we employ, particularly that for RPFM, should prove useful in delineating the relative importance of cues for modulation detection, including those that are stochastic.

II. METHOD

A. Stimuli

All stimuli were 500 ms in duration and were windowed with 5-ms raised-cosine on-off ramps. The three stimulus intervals of each trial were separated by 300 ms of silence.

The carrier of all signals was a 4000-Hz pure tone, with a level of 70 dB SPL prior to modulation, and modulation rates were 5, 20, and 40 Hz, in separate conditions. The SAM, QFM, and SFM signals were:

$$x_{SAM}(t) = [1 + m \cos(2\pi f_m t)] \cos(2\pi f_c t), \quad (1)$$

$$x_{QFM}(t) = \cos(2\pi f_c t) + (\Delta f / 2f_m) \{ \cos[2\pi(f_c - f_m)t] - \cos[2\pi(f_c + f_m)t] \}, \quad (2)$$

$$x_{SFM}(t) = \cos[2\pi f_c t + (\Delta f / f_m) \sin(2\pi f_m t)], \quad (3)$$

where f_c and f_m are the carrier and modulation frequencies in Hz, respectively, m is the modulation depth of the SAM signal, and Δf is the maximum instantaneous frequency excursion from the carrier frequency for the SFM signal. The quantity $(\Delta f / f_m)$ is β , the modulation index for SFM.

The RPFM signals were produced by first generating an SFM signal [Eq. (3)], then computing the fast Fourier transform (FFT) of the resulting signal. The phases were randomized by randomly choosing new phases for each component from a rectangular distribution over the range $0-2\pi$. Finally, the RPFM signal was reconstructed by computing an inverse FFT of the resulting spectrum made up of the original SFM magnitude spectrum and randomized phase spectrum. To compensate for possible changes in loudness due to modulation, the rms-amplitude of the modulated signal was set equal to that of the unmodulated stimulus. For the RPFM signals, a new set of random starting phases was generated on each trial.

All stimuli were generated digitally in Matlab and converted to analog signals at a 44.1 kHz sampling rate using a PC equipped with a 24-bit sound card (Echo Audio Gina). Stimuli were presented monaurally to the left ear over Sony MDR-V6 stereo headphones to listeners seated in an IAC sound-attenuating chamber.

B. Procedure

All thresholds were measured using a three-interval, three-alternative forced-choice paradigm, with a modulated stimulus presented in the signal interval and unmodulated stimuli presented in the remaining two intervals. Modulation depth was varied using a 3-down, 1-up adaptive procedure that tracked to the 79.4-percent correct point ($d' \approx 1.62$) on the psychometric function and for which chance performance is 33 percent correct.

SAM-detection thresholds were measured by varying the modulation depth and QFM-detection thresholds were measured by varying the modulation index, both in dB-like quantities ($20 \log m$ and $20 \log \beta$ for SAM and QFM, respectively). The step size was initially set to 2 dB and was reduced to 1 dB after the first four reversals. SFM-detection thresholds were measured by varying the maximum frequency excursion, Δf , in geometric steps by multiplying/dividing Δf by 1.32 for the first four reversals and 1.15 for subsequent reversals. For the detection of RPFM, Δf was varied geometrically in the same way as for SFM, with the Δf of Eq. (3) varied prior to randomization of the component phases. In all cases, each adaptive run continued until twelve

FIG. 1. Modulation-detection thresholds as a function of modulation frequency averaged across four listeners. The top-left panel shows detection thresholds for SFM, RPFM, QFM, and SAM plotted in terms of the level of the first spectral sideband adjacent to the 70 dB carrier. In the bottom-left panel, all but the SAM thresholds are plotted in terms of the total frequency excursion (TFE) for SFM and QFM, and that estimated for 1000 RPFM stimuli generated at the threshold level for each individual listener and averaged. In the bottom-right panel, SAM detection thresholds are plotted in terms of modulation depth, and the mean effective modulation depth (based on the maximum and minimum amplitudes of the Hilbert envelope) is plotted for the QFM and for 1000 RPFM stimuli. The lighter circles represent data from Edwards and Viemeister (1994) measured with a 1-kHz carrier. Error bars for the darker symbols represent standard errors.

reversals were obtained, with the mean of the final eight reversals taken as threshold for that run. The arithmetic mean was computed for SAM and QFM thresholds and the geometric mean for SFM thresholds. The mean of six threshold estimates was taken as the final threshold estimate for each condition. The modulation conditions were run in a pseudo-random order chosen by the experimenters.

Subjects initiated each run and made responses by pressing keys on the PC keyboard. Each interval was marked with a visual marker on the PC monitor and subjects received visual correct-answer feedback after each trial.

C. Subjects

Two of the four subjects were the second and third authors. The remaining subjects were female undergraduate students from the University of Minnesota who were paid to participate. All listeners had pure-tone thresholds of 15 dB HL or better at octave frequencies from 250–8000 Hz. All subjects had previous experience in other psychoacoustical tasks and were trained to asymptotic performance in the present experiment.

III. RESULTS AND DISCUSSION

The data for the four listeners were very similar in form and are well represented by the means, which are plotted in Fig. 1 with darker lines and symbols. The top-left panel of Fig. 1 shows the mean detection thresholds for all conditions plotted in terms of the level in dB SPL of the modulation sideband that is closest in frequency to the carrier (which for these stimuli at the measured threshold levels was the sideband with the highest level as well). The bottom-left panel of Fig. 1 shows the SFM, QFM, and RPFM thresholds expressed in terms of the range of the instantaneous frequency

of the modulated waveform, the total frequency excursion (TFE). This measure of the frequency fluctuations is somewhat arbitrary and was chosen for its intuitive appeal. For QFM and SFM the TFE is equal to $2\Delta f$ in Eqs. (2) and (3). An analysis using the standard deviation of the frequency fluctuations yielded a pattern of results similar to those shown in the bottom left panel of Fig. 1.

For RPFM, the TFE depends upon the component starting phases and thus varies across stimuli. In the bottom-left panel of Fig. 1, the TFE was estimated by generating 1000 stimuli with the threshold sideband level for each individual listener, computing the TFE for each sample, and then averaging the TFE across all 1000 samples and the four listeners. (The instantaneous frequency function was computed as the time derivative of the instantaneous phase of the analytic signal obtained using the Hilbert transform of the signal. The TFE is the difference between the maximum and minimum frequencies for a given sample.) It can be seen from the two left panels of Fig. 1 that the RPFM thresholds are substantially lower than those for SFM.

In the bottom-right panel of Fig. 1, SAM detection thresholds are shown along with the effective AM depth of the QFM stimuli and RPFM stimuli, both plotted in dB ($20 \log m$). The effective AM depth of the QFM stimuli was computed as $20 \log[(\max - \min)/(\max + \min)]$, where $\max$ and $\min$ are the maximum and minimum amplitudes of the Hilbert envelope at threshold. The effective AM depth of the RPFM stimuli in dB was estimated by computing $20 \log[(\max - \min)/(\max + \min)]$ for the 1000 stimuli generated as described above, where $\max$ and $\min$ are the maximum and minimum amplitudes of the Hilbert envelope of each stimulus. Those 1000 estimates of effective modulation depth were averaged in dB and are plotted in the bottom-right panel of Fig. 1. The effective modulation depths of the RPFM stimuli at threshold are in reasonably good agreement with the SAM detection thresholds across all modulation frequencies tested. This suggests that the primary cue for detecting RPFM is the envelope fluctuations inherent in the stimuli. Further supporting this is the indication that the changes in instantaneous frequency at RPFM threshold are well below the threshold for detecting the frequency changes in SFM and in QFM (bottom-left panel of Fig. 1). Apparently the frequency changes in RPFM are too rapid, too brief, or too irregular to provide a cue that is useful relative to that provided by the envelope fluctuations.

It is possible that the modulation was not being detected at the cochlear place corresponding to the carrier frequency, i.e., the observer may have been listening “off frequency” and using information from frequency regions remote from the carrier frequency. To assess this we included a condition in which a highpass noise (6-kHz cutoff, 45 dB spectrum level) was used to limit useful information above the carrier frequency (see Viemeister, 1988 regarding the “near miss” and AM detection). The noise increased the AM-depth thresholds for RPFM, SAM, and QFM similarly and so the pattern of results was nearly identical to that shown in the lower right panel of Fig. 1. This suggests that our results also

characterize “within-channel” processing and not some peculiarity resulting from off frequency or multi-channel listening.

Additionally, the data shown in Fig. 1 for QFM suggest that at a modulation frequency of 5-Hz QFM is detected on the basis of the external envelope fluctuations: The AM modulation depth of QFM at threshold is essentially identical to that for SAM (bottom-right panel). At the higher modulation frequencies, however, the AM modulation depth of the input at threshold for QFM decreases and the modulation depths are below those for detection of SAM. This suggests that at these modulation frequencies the QFM is *not* being detected on the basis of the external envelope fluctuations. At these frequencies (and at 5 Hz) the sideband levels (upper left panel) for SFM and QFM overlap perhaps indicating detection of sidebands, i.e., that the sidebands are resolved and mediate the detection of modulation. This seems unlikely considering the relatively low modulation frequencies and the 4-kHz carrier frequency where the bandwidth of the auditory filter is large relative to that of the stimuli. Also, the fact that sideband levels are lower for RPFM than for SFM, which have the same nominal amplitude spectra, further indicates that simple sideband resolution does not mediate modulation detection.

There are several plausible alternative explanations for the similarity of the threshold TFE’s for SFM and QFM at 20 and 40 Hz. It is possible that tonotopic changes produced by the changes in instantaneous frequency mediate detection. Another possibility is FM-to-AM conversion (Zwicker, 1956, 1962). Even though the auditory filter is broad at the 4-kHz carrier frequency the listener may be basing decisions on the responses from a filter tuned to a frequency slightly below the carrier frequency (see Kohlrausch *et al.*, 2000). Using the measured TFE for SFM detection and the AM modulation threshold for SAM, an attenuation rate of 177 dB/oct near 4 kHz would produce an FM-to-AM conversion consistent with the measured AM modulation threshold. This attenuation rate is reasonable based on psychophysical and physiological tuning curves (Moore, 1978; Robles *et al.*, 1986). Finally, a cue based on temporal coding of the stimulus fine structure seems unlikely because the carrier frequency is near the putative limit for phase locking (Weiss and Rose, 1988).

Also shown in Fig. 1 are data from Edwards and Viemeister (1994) using a 1-kHz carrier for conditions that are similar to those of the present experiment (lighter gray circles). (Their thresholds have been adjusted to correspond to the 79.4 percent-correct point being estimated in the present experiment. The adjustments were based on reconstructed psychometric functions and generally are minor.) Although there are substantial differences in thresholds, especially at the higher modulation frequencies, the general trends are similar for the 1- and 4-kHz carriers. A notable exception is the set of TFE data for QFM (bottom-left panel). At the lower modulation frequencies (4 and 5 Hz) the TFE thresholds for QFM are in close agreement for the two carrier frequencies. They are also in agreement for thresholds expressed as AM depth (bottom-right panel). As for the 4-kHz data this suggests that at these modulation frequen-

cies, the QFM is being detected on the basis of external envelope fluctuations. Furthermore, the TFE thresholds for QFM and SFM are similar for the 1-kHz carrier. This suggests that the SFM is also being detected based on envelope fluctuations produced by FM-to-AM conversion at least for 4-Hz modulation. This contrasts with the account of Moore and Sek (1996) who propose, based on data from experiments using mixed modulation (AM and FM), that for lower carrier frequencies and modulation frequencies below 10 Hz, SFM is detected by changes in the phase locking related to the changes in instantaneous frequency. Our suggestion is based on the possibly coincidental agreement in thresholds for QFM, SFM, and SAM and thus is relatively weak. The Moore and Sek suggestion relies on data obtained using mixed modulation, a situation that may involve interactions more complex than simple linear combination of envelopes. Unfortunately, there appear to be no auditory nerve recordings that are directly relevant to this issue.

In summary, the present data indicate that the large difference between the detectability of modulation of SFM and RPFM is primarily due to the external envelope fluctuations that are present in the RPFM. Since the amplitude spectra of these signals are identical, the difference cannot be due to spectral differences. Although there are periodic changes in instantaneous frequency in RPFM, these changes appear to provide a much less effective cue than those provided by the envelope fluctuations at least at the high carrier frequency used in these experiments. At the lowest modulation frequency (5 Hz) the threshold depth for SAM and that computed for QFM and RPFM, expressed as $20 \log m$, are essentially identical, suggesting a common mechanism that is based on external envelope fluctuations. At higher modulation frequencies the QFM thresholds in terms of TFE increase and coincide with those for SFM. At these modulation frequencies the computed AM depth for QFM is well below that for SAM, indicating that *external* envelope fluctuations for QFM are essentially undetectable. At these modulation frequencies, the similarity between the thresholds for SFM and QFM suggest a common mechanism, either FM-to-AM conversion or multi-channel place encoding. Because of the

high carrier frequency a temporal code based on fine structure is unlikely. Finally, comparison with previously published data using a 1-kHz carrier indicates similar general trends with the exception that the 1-kHz TFE thresholds for QFM do not increase with modulation frequency. The TFE threshold for QFM with a 4-kHz carrier at the lowest modulation frequency is very similar to the thresholds for QFM and SFM with a 1-kHz carrier. The similarity between these three thresholds and the similarity of the modulation thresholds for SAM with 1- and 4-kHz carriers suggests a common mechanism, one based on envelope fluctuations perhaps generated by auditory filtering.

ACKNOWLEDGMENTS

This work was supported by Research Grant No. R01 DC 00683 from the National Institute on Deafness and Communication Disorders, National Institutes of Health.

- Edwards, B. W., and Viemeister, N. F. (1994). "Psychoacoustic equivalence of frequency modulation and quasi-frequency modulation," *J. Acoust. Soc. Am.* **95**, 1510–1513.
- Kohlrausch, A., Fassel, R., and Dau, T. (2000). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," *J. Acoust. Soc. Am.* **108**, 723–734.
- Moore, B. C. J. (1978). "Psychophysical tuning curves measured in simultaneous and forward masking," *J. Acoust. Soc. Am.* **63**, 524–532.
- Moore, B. C. J., and Sek, A. (1996). "Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking," *J. Acoust. Soc. Am.* **100**, 2320–2331.
- Robles, L., Ruggero, M. A., and Rich, N. C. (1986). "Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases," *J. Acoust. Soc. Am.* **80**, 1364–1374.
- Viemeister, N. F. (1988). "Psychophysical aspects of auditory intensity coding," in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 213–241.
- Weiss, T. F., and Rose, C. (1988). "A comparison of synchronization filters in different auditory receptor organs," *Hear. Res.* **33**, 175–179.
- Zwicker, E. (1956). "Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs (The elementary bases for the determination of the information capacity of hearing)," *Acustica* **6**, 365–381.
- Zwicker, E. (1962). "Direct comparison between the sensations produced by frequency modulation and amplitude modulation," *J. Acoust. Soc. Am.* **34**, 1425–1430.