

**Viemeister, N.F., Rickert, M., Law, M., and Stellmack, M.A. (2002).
Psychophysical and physiological aspects of auditory temporal processing.
In Tranebjaerg, L., Christen-Dalsgaard, J., Andersen, T., and Poulsen, T.
(Eds.) Genetics and the Function of the Auditory System. Proceedings of
the 19th Danavox Symposium, Holmens Trykkeri, Denmark, pp. 273-291.**

PSYCHOPHYSICAL AND PHYSIOLOGICAL ASPECTS OF AUDITORY TEMPORAL PROCESSING

Neal F. Viemeister, Martin Rickert, Martin Law, and Mark Stellmack,
University of Minnesota

During the past two decades there has been an increasing interest in how the auditory system extracts information from dynamically varying sounds, sounds whose amplitude and/or frequency vary over time. The importance of such information has been clearly demonstrated in speech perception, sound localization, and in many basic aspects of monaural hearing. This chapter reviews recent experimental and theoretical findings that provide insights into the basic mechanisms involved in temporal processing, specifically those involved in monaural envelope processing. One area of interest is the possibility that the auditory system is “tuned” for envelope periodicity, i.e., that there are filters selectively tuned to the modulation frequency of amplitude modulation (AM). A powerful and provocative theoretical account has recently been proposed based on this notion. New psychophysical data relevant to this issue are discussed and a psychophysically-oriented approach for analyzing physiological data on envelope processing is presented. The second area to be considered is the role of timing information, “phase locking”, in monaural hearing. Several contemporary models have proposed that fine-structure timing information in the auditory nerve is fundamental in auditory coding. We present psychophysical data on octave phase effects that can be explained qualitatively by a simple model that does not involve the use of fine-structure information. This suggests that except for relatively slow envelope fluctuations, timing information may play a negligible role in monaural hearing.

Envelope processing: psychophysical aspects

There is a wealth of valuable psychophysical research on various aspects of monaural envelope perception. This includes work on Modulation Detection Interference (see Sheft and Yost, 1997), Comodulation Masking Release (see Buss, *et al.*, 1998), modulation depth discrimination (see Wakefield and Viemeister, 1990), modulation rate discrimination (see Lee, 1994) and pitch perception (see Burns and Viemeister, 1981), the relation between AM and FM detection (see Edwards and Viemeister, 1994; Moore and Sek, 1994), modulation rate uncertainty (Wright and Dai, 1998), and modulation detection by hearing-impaired listeners (see Bacon and Gleitman, 1992) and by cochlear-implant users (see Busby, *et al.*, 1993). A review of this important research is beyond the scope of this chapter and so the focus will be on a fairly recent development, modulation filtering, which may provide a unifying theoretical framework for interpreting the data on envelope perception.

The notion that the auditory system may be selectively tuned for modulation, both for AM and FM, appears to have been first suggested as an explanation for the results of experiments on selective adaptation, specifically, the finding that exposure to a given modulation frequency will raise the threshold for detecting modulation at that frequency but not for modulation frequencies far removed from the adapting frequency (Green and Kay, 1974). It was proposed that there are “channels” broadly tuned to modulation frequency. Another explanation, one that does not involve such channels, is that the threshold elevation and frequency selectivity results from uncertainty produced by the adaptor on the parameters of the signal (Wakefield and Viemeister, 1984). Although it has not been developed for AM detection, this “cognitive” explanation may provide an alternative to hard-wired frequency-selective channels.

A more compelling case for modulation filters for AM is based on results from experiments on modulation masking where it has been shown that modulation detection of a signal modulation frequency can be affected by masker modulation, a situation similar to that seen in traditional spectral (versus envelope) masking (Bacon and Grantham, 1989; Houtgast, 1989). Modulation filters are analogous to critical bands or auditory filters and have been proposed to account for the frequency selectivity shown in modulation masking. The most complete account is that of Dau and colleagues (Dau, *et al.*, 1997a; Dau, *et al.*, 1997b; Ewert and Dau, 2000). The model has gone through several refinements but its essential structure, a “modulation filterbank”, remains largely unchanged. The model provides a compelling account of many phenomena of modulation detection and has helped elucidate basic mechanisms that may be involved in envelope processing.

A particularly interesting aspect of the model is that it suggests that the Temporal Modulation Transfer Function (TMTF; Viemeister, 1977) obtained with broadband noise carriers may not reflect auditory sluggishness and temporal resolution as has been typically assumed (Viemeister, 1979). The argument is that inherent fluctuations in the noise carriers mask the imposed sinusoidal AM and that because the bandwidth of modulation filters increases at higher modulation frequencies, the amount of masking increases as the modulation frequency of the signal increases. Thus, modulation thresholds at higher modulation frequencies do not reflect “pure” envelope following and the cutoff frequency of the typical lowpass TMTF, an index of temporal resolution, is lower than the true cutoff, i.e., temporal resolution is better than is indicated by the TMTF measured with broadband noise carriers. Supporting this are experiments using high-frequency sinusoidal carriers (Kohlrausch, *et al.*, 2000) and low-noise noise carriers that have been constructed to reduce the envelope fluctuations inherent in Gaussian noise (Dau, *et al.*, 1999). The data shown in Fig. 1 illustrate these points: the TMTFs for sinusoidal carriers (circles) and for low-noise noise (triangles) have a higher cutoff frequency than does the TMTF for broadband noise carriers (squares). However, TMTFs with sinusoidal carriers, and more generally with bandlimited carriers, should be interpreted cautiously because modulation detection may not be based solely on envelope cues. In addition to the problem of sideband resolution at high modulation

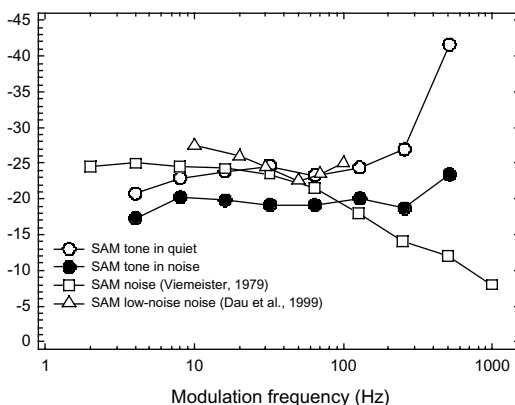


Fig. 1. Psychophysical TMTFs for various carrier conditions. The y-axis is the threshold modulation depth in dB for detecting sinusoidal AM (SAM) and is plotted as a transfer function with thresholds increasing downward.

frequencies (which likely accounts for the decreased thresholds above 200 Hz), there is the potential for detection based on distortion products (Wiegand and Patterson, 1999). There also are uncertainties regarding the role of off-frequency listening: for example, adding a highpass noise above the carrier frequency can raise modulation thresholds for sinusoidal carriers. To address these issues Rickert and Viemeister (1998) compared TMTFs obtained with a 2 kHz carrier at 62 dB SPL presented either in quiet (open circles in Fig. 1) or with a 15 dB spectrum level broadband noise masker (filled circles) that should be sufficient to mask distortion products and reduce off-frequency listening. The noise masker raised modulation thresholds but the TMTF is similar to that obtained in quiet except at the highest modulation frequencies, where sideband detection is involved. The conclusion is that the “flatness” of the TMTF does not result from detection of distortion products. Overall, these data suggest that temporal resolution is even better than that indicated by the TMTF obtained with broadband noise carriers and that the account of Dau *et al.* based upon modulation filters and modulation masking is probably correct.

Although the model proposed by Dau *et al.* provides a compelling account of these and other modulation detection data, direct experimental evaluation of the critical aspect of the model, modulation filters, has proved elusive and inconclusive. Several experiments have used complex modulators or maskers (Lorenzi, *et al.*, 1999; Moore, *et al.*, 1999; Sheft and Yost, 2001). There recently have been examinations of phase effects (Strickland and Viemeister, 1996; Moore and Sek, 2000) showing that modulation thresholds depend, sometimes substantially, on the phase relationship between sinusoidal modulation maskers and signal. According to the Dau *et al.* model, two components spaced an octave apart generally should be resolved by the modulation filters and so no substantial phase effects are predicted. A modification of the model by including envelope compression before the modulation filters appears to account for the data for signals an octave above the masker but not for signals an octave below (Moore and Sek, 2000; Verhey, *et al.*, 2001). An alternative, fundamentally different, approach is to use the traditional single-channel lowpass filter (versus a modulation filterbank) with decision statistics that capture local temporal features of the envelope (Strickland and Viemeister, 1996; Lorenzi, *et al.*, 1999). To date, this has only been moderately successful in accounting for phase effects. Whether such an approach can account for the frequency selectivity observed in modulation masking is unclear.

Envelope processing: physiological aspects

There are extensive physiological data on AM processing that provide insight into the issue of modulation tuning and, more generally, into how the auditory

system processes temporal information. Briefly, TMTFs based on envelope synchrony show a lowpass characteristic in the auditory nerve (Joris and Yin, 1992) and generally show tuning by the brainstem level and a decreased ability to synchronize to high modulation frequencies (see Krishna and Semple, 2000). The range of “best modulation frequencies” (BMFs) of cortical units appears to be limited: of the approximately 40-60% that are tuned, almost all have BMFs below about 100 Hz and a mean BMF of approximately 15 Hz (Schreiner and Urbas, 1988; Bieser and Müller-Preus, 1996).

Tuning in the cortex and IC has been shown for envelope synchrony and also for overall firing rate but the relationship between the TMTFs based on these measures is unclear, as is the role of inhibition in the rate-based response (Kuwada and Batra, 1999). It has been shown that changes in carrier level can change the TMTF from lowpass to bandpass and affect the BMF (Frisina, *et al.*, 1990a, 1990b; Rhode and Greenberg, 1994; Krishna and Semple, 2000). The effects of anesthesia and perhaps of attentiveness are a concern especially if the data are to be related to the psychophysical data. Of particular concern in this regard is the considerable discrepancy between the psychophysical and the cortical data on the upper limit for detection (2-4 kHz for 100% AM, see Fig. 1) and for melodic pitch perception (approximately 800 Hz; see Burns and Viemeister, 1976). Apparently, there are few, if any, cortical units that respond to these modulation frequencies.

In comparing the psychophysical and physiological data, there is an understandable tendency to simply compare functional dependencies such as TMTFs and draw inferences about the possible physiological basis for the psychophysical data. A potential problem with TMTFs is that in the physiological experiments the input modulation depth is held constant and some measure of response, usually a synchronization coefficient, is plotted as a function of modulation frequency. In psychophysical experiments, however, the modulation depth necessary for a constant response (“threshold”) typically is the dependent variable. A fundamental issue is the appropriate measure of response to compare with the psychophysical data. Another important issue is whether it is appropriate to compare a constant-input situation to that in which the output is presumably constant.

In some domains, such as intensity coding, there have been attempts to more directly relate the psychophysical and physiological data by simulating psychophysical experiments (e.g., Viemeister, 1988). To our knowledge this has not been attempted with AM and so we present a preliminary analysis of AM detection by a single unit in the MGB of an unanesthetized rabbit (the recording

procedures were similar to those described in (Kuwada and Batra, 1999). A two-interval, forced-choice (2IFC) psychophysical task was simulated in which one of the two intervals contained sinusoidal AM; the other contained only the 3093-Hz carrier at 40 dB SPL. The task for the neural observer was to choose the interval containing the AM. During each interval the decision statistic (e.g. a synchronization coefficient) was computed; the decision rule was to choose the interval for which the statistic was larger. A different 400-ms sample from the spike record for AM and for carrier alone was used on each trial. The percentage of correct responses based on at least 40 trials was then computed and converted to another measure of performance, d' . Neurometric functions ($\log d'$ vs. $20 \log m$) were constructed and from these the threshold modulation depth ($20 \log m$) necessary for $\log d' = 0$ (76% correct) was estimated. These thresholds are directly comparable to those measured behaviorally in that a similar task is being simulated, the performance measure is the same, and in both situations performance is being limited by variability in the decision statistic. The major unknown is the decision statistic ("decision variable", DV) used by the human observer, i.e. the perceptual attribute used by the human in making decisions about the presence/absence of AM. An important goal of this type of analysis, which is based on the concepts of Signal Detection Theory (see Green and Swets, 1966), is to find neural DVs that best describe the human data. The primary DVs examined were based on FFTs of the spike record during the 400-ms observation interval. One, total spike count, is proportional to the DC component in the FFT. The second was the amplitude of the component at the modulation frequency, a_{mod} . The third was this amplitude normalized by the DC component, a_{mod}/DC . This statistic is similar to the envelope-locked synchronization coefficient typically computed (over much longer averaging times) in studies of neural responses to AM. The fourth was the amplitude of the maximum non-DC component, a_{max} .

Note that the second and third DVs, those based on the component at the modulation frequency, assume, in effect, that the observer has precise knowledge of that frequency. One way to achieve this would be to have very sharply tuned modulation filters at some later stage of processing. This seems unrealistic. These DVs are considered here because they likely will produce "ideal" performance and because they are closely related to the synchronization coefficient. Neither spike count nor maximum non-DC requires knowledge of the modulation frequency.

Fig. 2 shows TMTFs based on three of the four DVs together with the human TMTF previously shown in Fig. 1 (open circles). The TMTF based on a_{mod} is similar to that for a_{mod}/DC and is not shown. The TMTF for a_{mod}/DC (open

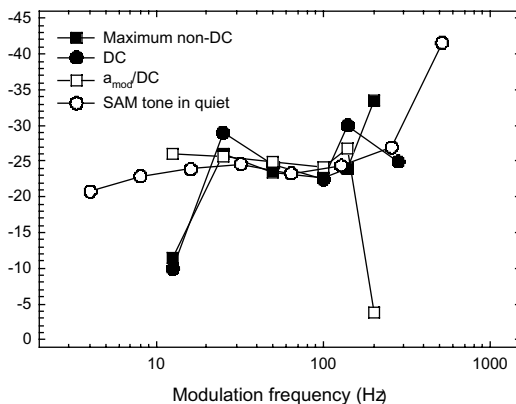


Fig. 2. Physiological TMTFs for various decision statistics. The open circles are human psychophysical data replotted from Fig. 1.

squares) shows essentially constant thresholds to approximately 140 Hz and an abrupt increase in threshold at 200 Hz (thresholds could not be estimated at higher modulation frequencies). The form of this TMTF is similar to that of the traditional synchronization function measured for this unit at small modulation depths (2% and 5%); however, the synchronization function shows a slight highpass segment, and therefore tuning, up to 100 Hz.

The similarity between the human thresholds and those based on a_{mod}/DC below 200 Hz is probably fortuitous and should be interpreted cautiously. Nevertheless, the similarity indicates that with ideal, or nearly ideal, processing a single unit in the MGB can show envelope sensitivity comparable to that shown perceptually.

The TMTFs based on DC (filled circles) and on a_{max} (filled squares) are different from those based on a_{mod} . Thresholds are considerably elevated at the lowest modulation frequency and are considerably lower at the highest frequency. The low threshold for a_{mod} at 200 Hz may reflect volleying, i.e., the strongest periodicity in the neural response may be for frequencies that are subharmonics of the modulation frequency.

Data from only one unit were suitable for analysis: only for this unit were recordings made for sufficiently small modulation depths, modulation depths for which estimated detection performance was less than perfect. Unfortunately, it appears that almost all of the available physiological data on responses to AM have been obtained with large modulation depths and are therefore unsuitable for this type of analysis. Clearly, the present analysis is far too limited to permit generalizations about neural AM processing. It seems safe to conclude, however, that ideal unit performance for AM detection can be comparable to that shown perceptually and, more importantly that the form of the TMTF can depend on the response characteristic, the DV, being considered.

Fine-structure timing

The role of fine-structure timing (vs. spectral-place and envelope) information in audition is an enormous, fundamental issue that dates to, at least, the famous Helmholtz-Seebeck controversy about pitch perception (see Moore, 1997). The definitions of “fine structure” and “envelope” are somewhat ambiguous but essentially refer to relatively fast changes in instantaneous pressure and to relatively slow changes in amplitude, respectively (see Viemeister and Plack, 1993). A change in fine structure that produces a change in the amplitude spectrum can be coded in the auditory system by a change in place of stimulation and, at least for frequencies below 4-5 kHz, by a change in the temporal pattern of discharges in the auditory nerve (see Weiss and Rose, 1988). For convenience, we refer to the latter simply as “fine-structure timing” or “timing” information, terms that are somewhat more general, perhaps, than “phase locking” or “synchrony”.

The psychophysical data clearly indicate that timing information is used in binaural hearing for frequencies below about 1.2 kHz (see Durlach and Colburn, 1978). The role of such information in monaural hearing or, more generally, in perceptual situations other than sound localization is unclear. There is a vast psychophysical and physiological literature relevant to this topic, including that on frequency discrimination (e.g. Goldstein and Srulovic, 1977), pitch perception (Schouten, 1970), intensity coding (Carney, *et al.*, 1999), masking (Moore, 1975), and speech perception (Young and Sachs, 1979). A fair summary is that this contentious issue is still unresolved, and that the current trend is a return toward timing-based models, (e.g., Meddis and Hewitt, 1991; Patterson, *et al.*, 1995; de Cheveigne, 1999; Shamma and Klein, 2000). From a coding perspective, an especially interesting model is that proposed by Carney and collaborators (Carney, 1994; Carney, *et al.*, 1999). According to this provocative computational model, timing information, together with coincidence detection,

is the primary basis for peripheral auditory coding of monaural information. Given these recent theoretical developments, a re-examination of the psychophysical data relevant to this fundamental issue is appropriate.

The psychophysical data that appear to be most relevant to the issue of timing information are those, loosely grouped, on “monaural phase effects” (Craig and Jeffress, 1962; Goldstein, 1967; Raiford and Schubert, 1971). These experiments usually involve discrimination between waveforms that have identical power spectra but differ in their phase spectra and therefore in their fine structure. Such waveforms are easily produced and, in many cases, are readily discriminable. This suggests the use of fine structure information. Unfortunately, such interpretation is generally, perhaps always, incorrect. The major problem is that the phase manipulations can produce envelope and instantaneous frequency differences (e.g. AM vs. quasi-FM). It is clear that such differences can be readily perceptible and, indeed, that they are crucially important in auditory perception. Another problem is that the phase manipulations may alter auditory distortion products, such as the cubic difference tone, and thus may provide a cue for discrimination that is not based on fine structure.

Two-tone octave complexes offer an appealing solution to these problems and provide a useful tool for examining the role of fine structure information. With such stimuli there are no useful spectral or envelope cues and there are no “classical” distortion products that can produce significant phase effects. It is well known that a slight mistuning (e.g., 5 Hz) of the octave can produce beats. This can be considered a phase effect in that it is similar to a situation in which the two tones are in a perfect octave but where the phase relation varies continuously during the “beat” cycle. A mistuned octave and, more generally, mistuned consonances show periodic changes in the waveform. The fact that the beats of mistuned consonances are audible has been taken as evidence for the importance of temporal fine structure in monaural hearing (Plomp, 1967). We (Viemeister, *et al.*, 2001) have shown, however, that the beats of a mistuned octave can be perceived when the complex consists of 4000 Hz and 8005 Hz. At these frequencies the representation of fine structure is severely degraded in the auditory nerve. These data thus suggest that fine-structure information is not necessary to detect the beats of mistuned consonances.

An alternative, one that does not require fine-structure information, includes simple compression after linear combination of the components. More specifically, the model consists of bandpass (auditory) filtering, followed by rectification, followed by instantaneous compression (exponentiation to the power 0.3), followed by lowpass filtering. This model extends the linear

envelope detector by using compression and is similar to other recent models using compression (Kohlrausch and Sander, 1995; Carlyon and Datta, 1997; Oxenham and Moore, 1997). A linear envelope detector shows no periodic fluctuations at the “beat” frequency of a mistuned octave. With compression, however, there is a clear periodicity at the beat frequency and its effective modulation depth roughly agrees with the psychophysical data (Viemeister, *et al.*, 2001). The key notion is that the changing phase relation between the two components of the mistuned octave produces periodic changes in the effective level at some cochlear place where the two components can add and be subjected to compression.

To further explore the role of fine-structure timing, we (Law *et al.*, 2000) examined phase discrimination of perfect octave complexes, a situation analogous to “freezing” the mistuned octave at some point in time. Low frequencies were used in an attempt to ensure that strong timing information was available. In the first part of the study, detection thresholds for a 500-Hz tone were measured in the presence of a 250-Hz masker presented for 500 ms at 55, 70, and 80 dB SPL. Fig. 3 shows thresholds as a function of the phase relation between the components. These data are similar to those of Nelson and Bilger (1974) and of Clack *et al.*, (1972) in that they show a clear dependence of threshold on phase. Nelson and Bilger interpreted this dependence as resulting from changes in fine structure. However, the simple model outlined above also predicts a phase dependence because changes in phase will produce changes in effective level and thus changes in detection threshold.

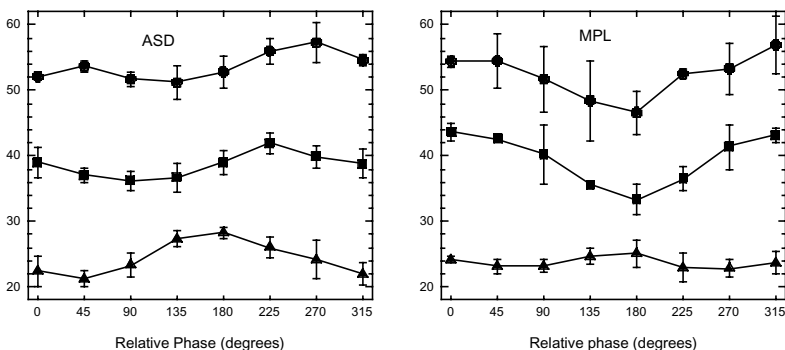


Fig. 3. Thresholds for a 500-Hz tone masked by a 250-Hz tone. The level of the masker was 55 (triangles), 70 (squares), or 80 dB SPL (diamonds). Data are shown for two subjects.

The major part of the present study examined the discriminability of phase differences in the octave complexes. Based on the detection data from the first part of the study (Fig. 3), phases and levels were chosen such that a phase difference would be highly discriminable under certain conditions. Specifically, in the non-signal interval of the two-cue-2IFC task the two components were in a 45-deg phase relation and the level of the 500-Hz component was 3 dB above the detection threshold (those shown in Fig. 3). In the signal interval the relative phase of the upper component was changed by 180 deg (to 225 deg) and percent correct responses [P(C)] was measured as a function of the relative level of the upper component.

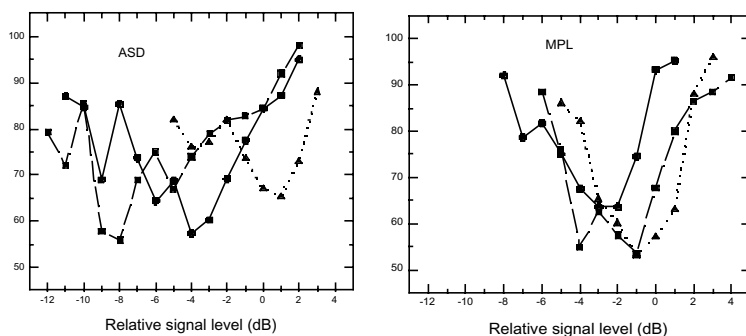


Fig. 4. Psychometric functions for detecting a 180 deg phase change between the components of an octave complex at 250 Hz. The x-axis is the level of the 500-Hz component in the signal interval relative to that in the non-signal interval. Data for the same two subjects for Fig. 3 are shown.

The rationale for this experiment was that if discrimination of phase differences were based on differences in effective level then it should be possible to compensate for these level differences by external level differences and thereby render this cue unusable. Fig. 4 shows the psychometric functions for the two subjects. The x-axis is the level of the 500-Hz component in the signal interval relative to that in the non-signal interval. The V-shape of these functions indicates that small changes in relative level can markedly affect the discrimination of the 180-deg phase change. Since fine-structure information is unlikely to be affected by such small changes in relative level, we conclude that phase discrimination near the minimum of the V was based primarily on changes

in effective level. Conditions in which performance degraded to chance ($P(C) = 0.5$) strongly suggest that the discrimination in these conditions was based on level cues. Where minimum performance did not go to chance, timing information may account for the remaining ability to discriminate the phase change. In principle, it should be possible to use such data to separate the relative contributions of timing and level information and to quantitatively predict performance when both are present. Unfortunately, it is unclear how to independently estimate these contributions experimentally and to assess how the information is combined.

In summary, the data on the beats of mistuned octaves, on phase effects in octave masking, and on phase discrimination of octave complexes are all compatible with a simple model in which envelope extraction occurs after a compressive nonlinearity. It appears that even at low frequencies, where fine-structure timing information is robust, the phase-induced level cue is dominant over timing information.

Concluding comments

We have attempted to highlight contemporary issues that are important for understanding auditory temporal processing. The possible role of tuning for AM, specifically the notion of modulation filters, needs further experimental examination. It is not clear, however, what line of psychophysical research may be most useful and, at present, no definitive series of experiments that will resolve this basic issue is apparent. Hampering this effort is that there is no compelling alternative model that describes the tuning seen psychophysically in modulation masking experiments but does not incorporate “hard-wired” modulation filters. The physiological data on AM processing do suggest tuning at central loci but there remain uncertainties about the appropriate response measures and, of course, whether the units that show tuning play a role in perception. To better relate the psychophysics to the physiology we considered an analysis scheme that simulates a perceptual task and which forces explicit assumptions about the basis for decisions in the perceptual task. The analysis was limited to only one MGB unit but illustrated the basic approach and, more importantly, clearly indicated that inferences about tuning depend on what decision variable is assumed.

The other general area of concern was the role of fine-structure timing information in monaural hearing. This recurring issue is fundamental for understanding auditory coding and is firmly at the intersection of physiology and psychophysics. Our re-examination of the role of timing information has been stimulated by recent physiologically-based models in which fine-structure timing plays a crucial role. The psychophysical data presented here are an attempt to address this issue experimentally. They suggest that certain monaural phase effects, effects that have been taken as demonstrating the importance of timing, can be explained simply by generalizing a familiar envelope detection model by incorporating a compressive nonlinearity. This raises the fundamental question as to whether fine-structure information plays any significant role in monaural hearing.

Acknowledgements

We thank Drs. Douglas Fitzpatrick and Shigyuki Kuwada for providing the MGB data used in the analysis and for stimulating discussions about the physiology of envelope processing. Thanks also to Dr. Magdalena Wojtczak for comments on the manuscript. Supported by NIH/NIDCD grant DC00683.

References

- Bacon, S.P. and Gleitman, R.M. (1992). "Modulation detection in subjects with relatively flat hearing losses," *J. Speech Hear. Res.* **35**, 642-653.
- Bacon, S.P. and Grantham, D.W. (1989). "Modulation masking: effects of modulation frequency, depth, and phase," *J. Acoust. Soc. Am.* **85**, 2575-2580.
- Bieser, A. and Müller-Preus, P. (1996). "Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds.," *Exp. Brain Res.* **108**, 273-284.
- Burns, E.M. and Viemeister, N.F. (1976). "Nonspectral pitch," *J. Acoust. Soc. Am.* **60**, 863-869.
- Burns, E.M. and Viemeister, N.F. (1981). "Played-again SAM: further observations on the pitch of amplitude-modulated noise," *J. Acoust. Soc. Am.* **70**, 1655-1660.
- Busby, P.A., Tong, y.C. and Clark, G.M. (1993). "The perception of temporal modulations by cochlear implant patients," *J. Acoust. Soc. Am.* **94**, 124-131.
- Buss, E., Hall III, J.W. and Grose, J.H. (1998). "Change in envelope beats as a possible cue in comodulation masking release (CMR)," *J. Acoust. Soc. Am.* **103**, 1592-1597.
- Carlyon, R.P. and Datta, A.J. (1997). "Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system," *J. Acoust. Soc. Am.* **101**, 3636-3647.
- Carney, L.H. (1994). "Spatiotemporal coding of sound level: Models for normal encoding and recruitment of loudness," *Hear. Res.* **76**, 31-44.
- Carney, L.H., Heinz, M.G. and Colburn, H.S. (1999). "Spatiotemporal coding of sound level: Quantifying the information provided by level-dependent phase cues," *Abstr. 22nd ARO Midwinter meeting* , 848.
- Clack, T.D., Erdreich, J. and Knighton, R.W. (1972). "Aural harmonics: the monaural phase effects at 1500 Hz, 2000 Hz, and 2500 Hz observed in tone-on-tone masking when $f/\text{sub } 1/=1000 \text{ Hz}$," *J. Acoust. Soc. Am.* **52**, 536-542.

Craig, J.H, Jeffress, L.A. (1962). "Effect of phase on the quality of two-component tone," J. Acoust. Soc. Am. **34**, 1752-1760.

Dau, T., Kollmeier, B. and Kohlrausch, A. (1997a). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," J. Acoust. Soc. Am. **102**, 2892-2905.

Dau, T., Kollmeier, B. and Kohlrausch, A. (1997b). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," J. Acoust. Soc. Am. **102**, 2906-2919.

Dau, T., Verhey, J. and Kohlrausch, A. (1999). "Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers," J. Acoust. Soc. Am. **106**, 2752-2760.

de Cheveigne, A. (1999). "Pitch shifts of mistuned partials: A time-domain model," J. Acoust. Soc. Am. **106**, 887-897.

Durlach, N.I. and Colburn, H.S. (1978). "Binaural Phenomena," in *Handbook of Perception, Vol. IV, Hearing*, edited by E.C. Carterette and M.P. Friedman (Academic Press, New York), pp. 365-466.

Edwards, B.W. and Viemeister, N.F. (1994). "Frequency modulation versus amplitude modulation discrimination: Evidence for a second frequency modulation encoding mechanism," J. Acoust. Soc. Am. **96**, 733-740.

Ewert, S.D. and Dau, T. (2000). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181-1196.

Frisina, R.D., Smith, R.L. and Chamberlain, S.C. (1990a). "Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement," Hearing Res. **44**, 99-122.

Frisina, R.D., Smith, R.L. and Chamberlain, S.C. (1990b). "Encoding of amplitude modulation in the gerbil cochlear nucleus. II. Possible neural mechanisms," Hearing Res. **44**, 123-142.

Goldstein, J.L. (1967). "Auditory spectral filtering and monaural phase perception," J. Acoust. Soc. Am. **41**, 458-479.

Goldstein, J.L. and Sruлович, P. (1977). "Auditory-nerve spike intervals as an adequate basis for aural frequency measurement," in *Psychophysics and Physiology of Hearing*, edited by E.F. Evans and J.P. Wilson (Academic Press, London), pp. 337-346.

Green, D.M. and Swets, J.A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, reprinted by R.E. Krieger, Peninsula Publishing, Los Altos, CA).

Green, G.G. and Kay, R.H. (1974). "Channels in the human auditory system concerned with the waveform of modulation present in amplitude- and frequency-modulated tones," *J. Physiol., London.* **241**, 29-30P.

Houtgast, T. (1989). "Frequency selectivity in amplitude-modulation detection," *J. Acoust. Soc. Am.* **85**, 1676-1680.

Joris, P.X. and Yin, T.C.T. (1992). "Responses to amplitude-modulated tones in the auditory nerve of the cat," *J. Acoust. Soc. Am.* **91**, 215-232.

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.

Kohlrausch, A. and Sander, A. (1995). "Phase effects in masking related to dispersion in the inner ear.II. Masking period patterns of short targets," *J. Acoust. Soc. Am.* **97**, 1817-1829.

Krishna, B.S. and Semple, M.N. (2000). "Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus," *J. Neurophysiol.* **84**, 255-273.

Kuwada, S. and Batra, R. (1999). "Coding of sound envelopes by inhibitory rebound in neurons of the superior olivary complex in the unanesthetized rabbit," *J. Neuroscience* **19**(6), 2273-2287.

Law, M.P., Stellmack, M.A. and Viemeister, N.F. (2000). "Monaural phase effects: Timing versus level cues," *J. Acoust. Soc. Am.* **107**, 2881 (Abstract).

Lee, J. (1994). "Amplitude modulation rate discrimination with sinusoidal carriers," *J. Acoust. Soc. Am.* **96**, 2140-2147.

Lorenzi, C., Berthommier, F. and Demany, L. (1999). "Discrimination of amplitude-modulation phase spectrum," *J. Acoust. Soc. Am.* **105**, 2987-2990.

Meddis, R. and Hewitt, M.J. (1991). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification," *J. Acoust. Soc. Am.* **89**, 2866-2882.

Moore, B.C.J. (1975). "Mechanisms of masking," *J. Acoust. Soc. Am.* **57**, 391-399.

Moore, B.C.J.M. (1997). "*An Introduction to the Psychology of Hearing*", 5th ed. (Academic Press, New York).

Moore, B.C.J. and Sek, A. (1994). "Discrimination of modulation type (amplitude modulation or frequency modulation) with and without background noise," *J. Acoust. Soc. Am.* **96**, 726-732.

Moore, B.C.J. and Sek, A. (2000). "Effects of relative phase and frequency spacing on the detection of three-component amplitude modulation," *J. Acoust. Soc. Am.* **108**, 2337-2344.

Moore, B.C.J., Sek, A. and Glasberg, B.R. (1999). "Modulation masking produced by beating modulators," *J. Acoust. Soc. Am.* **106**, 908-918.
Nelson, D.A. and Bilger, R.C. (1974). "Pure-tone octave masking in normal hearing listeners," *J. Sp. Hear. Res.* **17**, 223-251.

Oxenham, A.J. and Moore, B.C.J. (1997). "Modeling the effects of peripheral nonlinearity in listeners with normal and impaired hearing," in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Hillsdale, NJ), pp. 273-288.

Patterson, R.D., Allerhand, M.H. and Giguere, C. (1995). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.* **98**, 1890-1894.

Plomp, R. (1967). "Beats of mistuned consonances," *J. Acoust. Soc. Am.* **42**, 462-474.

Raiford, C.A. and Schubert, E.D. (1971). "Recognition of phase changes in octave complexes," *J. Acoust. Soc. Am.* **50**, 559-567.

Rhode, W.S. and Greenberg, S. (1994). "Encoding of amplitude modulation in the cochlear nucleus of the cat," *J. Neurophysiol.* **71**, 1797-1825.

Rickert, M.E. and Viemeister, N.F. (1998). "Temporal versus spectral cues in AM detection," *Proc. ICA/ASA98*, 825-826.

Schouten, J.F. (1970). "The residue revisited," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. Smoorenburg (A.W. Sijthoff, Leiden), pp. 41-58.

Schreiner, C.E. and Urbas, J.V. (1988). "Representation of amplitude modulation in the auditory cortex of the cat, II. Comparison between cortical fields.," *Hear Res* **32**, 49-63.

Shamma, S. and Klein, D. (2000). "The case of the missing pitch templates: How harmonic templates emerge in the early auditory system," *J. Acoust. Soc. Am.* **107**, 2631-2644.

Sheft, S. and Yost, W.A. (1997). "Modulation detection interference with two-component masker modulators," *J. Acoust. Soc. Am.* **102**, 1106-1112.

Sheft, S. and Yost, W.A. (2001). "AM detection with interrupted modulation," in *Physiological and psychophysical bases of auditory function*, edited by D.J. Breebart, A.J.M. Houtsma, A. Kohlrausch, V.F. Prijs, and R. Schoonhoven (Shaker, Maastricht), pp. 290-297.

Strickland, E.A. and Viemeister, N.F. (1996). "Cues for discrimination of envelopes," *J. Acoust. Soc. Am.* **99**, 3638-3646.

Verhey, J.L., Ewert, S. and Dau, T. (2001). "Modulation masking produced by masker modulation tone complexes," *J. Acoust. Soc. Am.* **109**, 2466 (Abstract).

Viemeister, N.F. (1977). "Temporal factors in audition: A systems analysis approach," in *Psychophysics and Physiology of Hearing*, edited by E.F. Evans and J.R. Wilson (Academic Press, London), pp. 419-428.

Viemeister, N.F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364-1380.

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.

Viemeister, N.F. and Plack, C. (1993). "Time analysis," in *Human Psychophysics*, edited by W. Yost, A. Popper, and R. Fay (Springer-Verlag, New York), Vol. 3.

Viemeister, N.F., Rickert, M. and Stellmack, M. (2001). "Beats of mistuned consonances: implications for auditory coding," in *Physiological and psychophysical bases of auditory function*, edited by D.J. Breebart, A.J.M. Houtsma, A. Kohlrausch, V.F. Prijs, and R. Schoonhoven (Shaker, Maastricht), pp. 113-120.

Wakefield, G.H. and Viemeister, N.F. (1984). "Selective adaptation to linear frequency-modulated sweeps: evidence for direction-specific FM channels?," *J. Acoust. Soc. Am.* **75**, 1588-1592.

Wakefield, G.H. and Viemeister, N.F. (1990). "Discrimination of modulation depth of sinusoidal amplitude modulation (SAM) noise," *J. Acoust. Soc. Am.* **88**, 1367-1373.

Weiss, T.F. and Rose, C. (1988). "A comparison of synchronization filters in different auditory receptor organs," *Hearing Res.* **33**, 175-179.

Wiegrebe, L. and Patterson, R.D. (1999). "Quantifying the distortion products generated by amplitude-modulated noise." *J. Acoust. Soc. Am.* **106**, 2709-2718.

Wright, B.A. and Dai, H. (1998). "Detection of sinusoidal amplitude modulation at unexpected rates," *J. Acoust. Soc. Am.* **104**, 2991-2996.

Young, E.D. and Sachs, M.B. (1979). "Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers," *J. Acoust. Soc. Am.* **66**, 1381-1403.