DTU Library

Perceptual interaction between carrier periodicity and amplitude modulation in broadband stimuli: A comparison of the autocorrelation and modulation-filterbank model

Stein, A.; Ewert, Stephan; Wiegrebe, L.

[Link back to DTU Orbit](#)

Perceptual interaction between carrier periodicity and amplitude modulation in broadband stimuli: A comparison of the autocorrelation and modulation-filterbank model

Alexandra Stein
Department Biologie II, Universität München, Großhadernerstrasse 2, 82152 Planegg-Martinsried, Germany

Stephan D. Ewert
Centre for Applied Hearing Research, Acoustic Technology, Ørsted DTU, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

Lutz Wiegrebe[a)]
Department Biologie II, Universität München, Großhadernerstrasse 2, 82152 Planegg-Martinsried, Germany

Recent temporal models of pitch and amplitude modulation perception converge on a relatively realistic implementation of cochlear processing followed by a temporal analysis of periodicity. However, for modulation perception, a modulation filterbank is applied whereas for pitch perception, autocorrelation is applied. Considering the large overlap in pitch and modulation perception, this is not parsimonious. Two experiments are presented to investigate the interaction between carrier periodicity, which produces strong pitch sensations, and envelope periodicity using broadband stimuli. Results show that in the presence of carrier periodicity, detection of amplitude modulation is impaired throughout the tested range (8–1000 Hz). On the contrary, detection of carrier periodicity in the presence of an additional amplitude modulation is impaired only for very low frequencies below the pitch range (<33 Hz). Predictions of a generic implementation of a modulation-filterbank model and an autocorrelation model are compared to the data. Both models were too insensitive to high-frequency envelope or carrier periodicity and to infra-pitch carrier periodicity. Additionally, both models simulated modulation detection quite well but underestimated the detrimental effect of carrier periodicity on modulation detection. It is suggested that a hybrid model consisting of bandpass envelope filters with a ripple in their passband may provide a functionally successful and physiologically plausible basis for a unified model of auditory periodicity extraction.
© 2005 Acoustical Society of America. [DOI: 10.1121/1.2011427]

I. INTRODUCTION

Within the last decades, many psychoacoustical models have converged to share relatively detailed implementations of auditory peripheral processing. Moreover, these models have been very successful in explaining a large variety of psychoacoustical phenomena, among them periodicity pitch and amplitude-modulation detection.

Even though it is unquestioned that envelope periodicity (such as sinusoidal amplitude modulation) imposed on broadband carriers is extracted in the time domain, the exact neural basis for the perception of harmonic complex sounds which elicit strong pitch sensations is still under debate. One line of research favors frequency domain models which evaluate the degree of neural excitation at harmonically spaced places along the auditory tonotopic axis (Cohen *et al.*,

1995; Goldstein, 1973; Terhardt, 1974). Time-domain models, in contrast, extract the pitch of harmonic sounds through an analysis of the temporal firing characteristics of the auditory nerve.

One of these time-domain models was introduced by Meddis and Hewitt (1991a, b). This model is based on Licklider's (1951) suggestion that the auditory system performs a process similar to autocorrelation in order to extract the periodicity of sounds. In addition, Meddis and Hewitt (1991a) suggested that the result of this autocorrelation is aggregated along the tonotopic frequency axis to produce a purely temporal representation of periodicity. This model accounts for various pitch perception phenomena, e.g., the pitch of the missing fundamental, the existence region, and the dominance region of pitch. Recent tests have also included the pitches of harmonic complexes with alternating phases and sensitivity to mistuned harmonics (Meddis and O'Mard,

a)Author to whom correspondence should be addressed; electronic mail: lutzw@lmu.de

1997). With some exceptions, (Carlyon, 1998), this model has been quite successful in predicting the results of many experiments on pitch perception.

A different model, designed to simulate the perception of amplitude modulation, was introduced by Dau *et al.* (1997a, b). It is based on the hypothesis that the envelope wave form undergoes a spectral decomposition similar to the decomposition of the wave form in the inner ear (Houtgast, 1989). This hypothesis is based on the demonstration of frequency selectivity in the envelope domain. Dau *et al.* (1997a) designed a modulation filterbank as a functional implementation of envelope-frequency selectivity. As in the above-mentioned pitch model, the modulation filterbank is preceded by a relatively detailed simulation of peripheral preprocessing stages. The modulation-filterbank model has been quite successful in simulating many aspects of amplitude-modulation perception, e.g., modulation detection for narrow-band noise carriers and envelope-frequency selectivity.

Even though both models aim to simulate different types of experiments, their basic design is very similar. The models share a gammatone filterbank and a functional simulation of mechano-neural transduction in inner hair cells. In both models, these stages are followed by a module which performs a temporal periodicity extraction which in Dau *et al.* (1997a, b) is implemented as a modulation filterbank and in Meddis and Hewitt (1991a, b) and Meddis and O'Mard (1997) is implemented as an autocorrelator. Even though the existence region of pitch and amplitude-modulation detection differs (Krumbholz *et al.*, 2000; Viemeister, 1979), there is still considerable overlap in the periodicity range between about 30 and 1000 Hz.

Most pronounced pitches are generated by stimuli with a periodic carrier which results in a harmonic power spectrum. Supposing that both carrier periodicity (CP) and amplitude modulation (AM) is extracted in the time domain, it is not parsimonious to assume two independent neural mechanisms of periodicity extraction. Moreover, there is already some evidence that carrier periodicity and modulation perception for broadband carrier are not independent. Burns and Viemeister (1976; 1981) demonstrated that sinusoidally amplitude modulated (SAM) noise can also produce a tonal sensation and convey melodic information. In SAM noise, the carrier is Gaussian noise (GN), which is aperiodic. Therefore, the pitch information can only be conveyed by the envelope.

This paper is concerned with the perceptual interaction between envelope and carrier periodicity in broadband stimuli. A typical stimulus with envelope periodicity is SAM noise. In contrast, iterated rippled noise (IRN) is a stimulus with pronounced carrier periodicity. Although this carrier periodicity is reflected, to some extent, in the IRN envelope, a series of psychoacoustical experiments have demonstrated that the perception of IRN is dominated by the carrier periodicity, not by the envelope periodicity (Wiegrebe and Patterson, 1999; Yost *et al.*, 1998). IRN is derived from GN. To generate this stimulus, a sample of GN is subjected to a delay-and-add process, which is iterated several times (Yost, 1996). This process results in a quasiperiodic stimulus. The magnitude spectrum of an IRN stimulus shows a ripple with peaks at the reciprocal of the delay and integer frequency multiples. As a consequence, the fundamental frequency (f0) of IRN corresponds to the reciprocal of the delay.

Thus, there are two broadband stimuli at hand, where the perception is dominated either by the envelope or by the carrier. Moreover, both stimuli allow for a continuous variation of the degree of periodicity. In SAM noise, this is done by varying the modulation depth of the modulator whereas in IRN, the degree of periodicity can be varied by changing the gain factor in the delay-and-add loop. When the gain factor is set to unity, the delay-and-add loop produces the maximum periodicity, determined by the number of iterations. A gain factor of zero results in an unchanged GN wave form at the output of the loop. In the current study the nomenclature $IRN(d, g, n)$ is used to specify the delay, d, in ms; the linear gain, g, and the number of iterations, n, respectively. To provide an easier comparison of carrier and envelope periodicity, the term fundamental frequency (f0) is used instead of delay for IRN.

An exemplary illustration of the recruited stimuli is shown in Fig. 1: The stimuli are SAM noise with a modulation frequency of 125 Hz and 0 dB (100%) modulation index, an $IRN(8, 1, 16)$ with a delay corresponding to 125 Hz, and a SAM IRN where the sinusoidal amplitude modulation was imposed on an IRN carrier instead of a Gaussian-noise carrier. Note that in this stimulus, the modulation frequency always equals f0. The illustration shows the time and frequency domain representations of exemplary versions of the stimulus wave forms (first and second rows) and their Hilbert envelopes (third and fourth rows). Both the audio- and the envelope spectrum of the stimuli with a carrier periodicity (IRN or SAM IRN) are harmonically structured. While the AM does not change the audio spectrum, it adds a distinct peak at the modulation frequency (indicated by the arrows) in the envelope spectrum. The highest component in the envelope spectrum is the dc component equivalent to the mean of the envelope wave form. It was fixed at 100 dB. Comparing the SAM peak height in the envelope spectrum with the magnitude of the spectral ripple of the IRN envelope spectrum, the overall envelope magnitude is much smaller for the IRN stimulus than for the stimuli with the SAM. As shown in the following, this feature is critical for the encoding of the IRN envelope spectrum with modulation filters having an appreciable dc sensitivity.

In the following, two psychoacoustic experiments are presented. The first experiment investigated to which extent AM detection was impaired by the presence of CP. The second experiment investigated to which extent CP detection was impaired by the presence of AM. In the second part of the paper, the predictions of an autocorrelation model and of a modulation-filterbank model are compared to the experimental data. Both models were tested on each set of obtained data. Finally, the basic differences between modulation filters and autocorrelation are highlighted and possible improvements of the models are discussed.

SAM Noise
Waveform

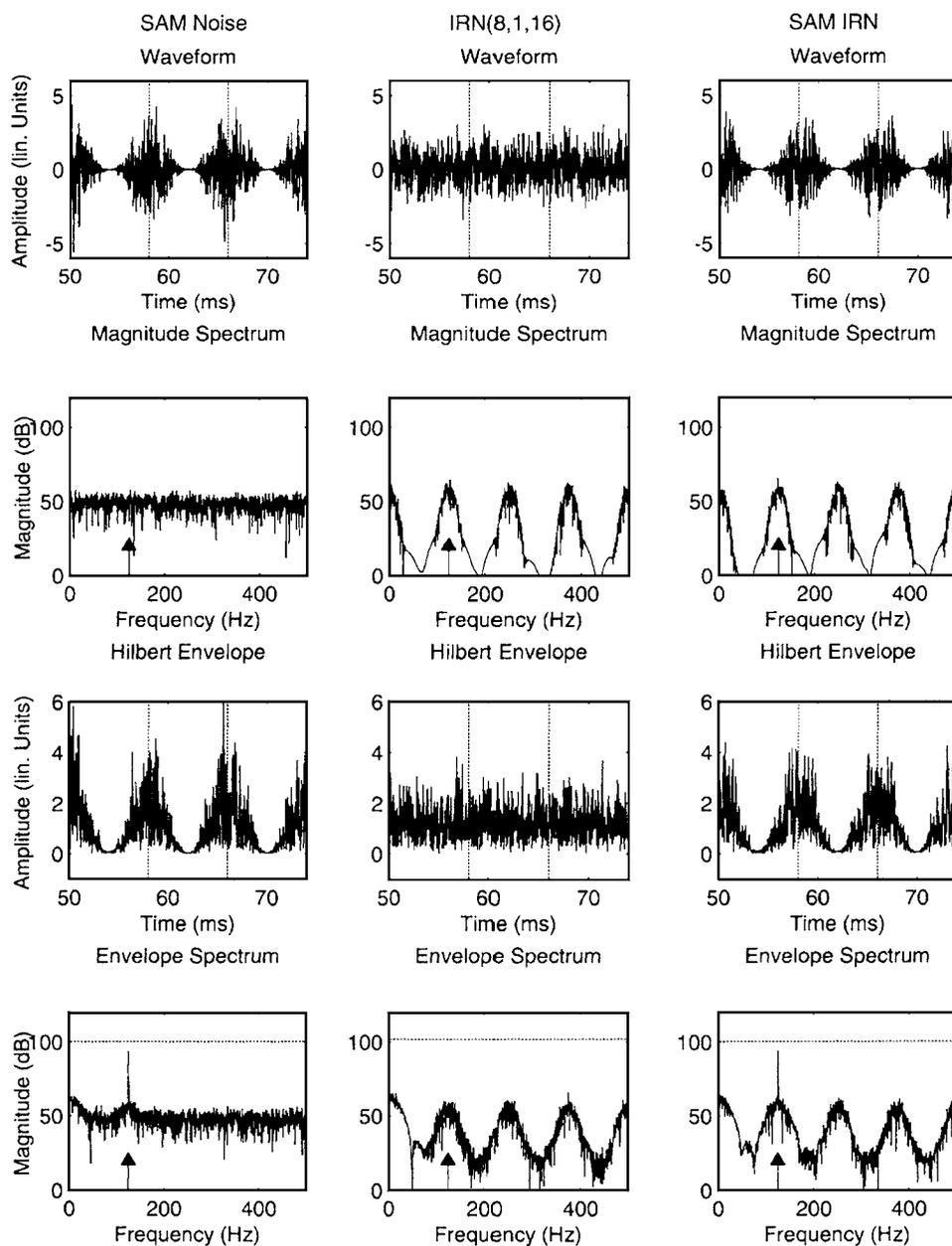IRN(8,1,16)
Waveform

SAM IRN
Waveform

FIG. 1. Examples of the stimuli used in the present study: Stimuli are either sinusoidally amplitude modulated noise (first column), iterated rippled noise (second column), or iterated rippled noise with an additional sinusoidal amplitude modulation superimposed (third column). The modulation frequency and the IRN f0 are 125 Hz (indicated by the arrows). The four rows show the stimulus wave form, the wave form magnitude spectrum, the Hilbert envelope, and the magnitude spectrum of the Hilbert envelope, respectively. Note that both the wave form and the envelope spectrum of IRN stimuli have a harmonic structure. The amplitude modulation does not change the wave form spectrum but it generates a distinct peak at the modulation frequency in the envelope spectrum. Due to the stochastic nature of the stimuli, the plots represent only sample realizations of the stimuli.

II. EXPERIMENTS

A. Experiment I: The effect of CP on AM detection

1. Rationale

This experiment examined the influence of CP on the detection of AM. Listeners were asked to detect a sinusoidal AM imposed on a broadband carrier. The carrier was either aperiodic (GN) or quasiperiodic (IRN). When an IRN carrier was used, the IRN f0 always matched the modulation frequency. Since the magnitude spectrum of IRN shows harmonic peaks at integer multiples of f0, the modulation sidebands fall on the peaks of the spectral ripples. This minimized the spectral cues available to the listeners.

2. Method

a. Stimuli and apparatus. The stimulus carrier was either GN or IRN. In the signal interval, a sinusoidal amplitude modulation was applied. Modulation frequencies ranged from 7.8125 to 1000 Hz in one-octave steps. The IRN carrier was generated in the add-same configuration (Yost, 1996). This add-same configuration was used for all IRN stimuli in this paper. The IRN gain was set to unity and 16 iterations were used. The duration of all stimuli was 500 ms including 20 ms, raised-cosine ramps. The stimuli were digitally generated on an IBM compatible PC at a sampling rate of 44.1 kHz and subsequently downloaded onto a TDT AP2 array processor. They were then converted to analog signals by a TDT DA 3-2 D/A converter. All stimuli were amplified by a Yamaha M35 2/4 Channel Power Amplifier followed by a passive end attenuation of 40 dB. The stimuli were presented diotically via AKG K240DF headphones. The average presentation level was 60 dB SPL. In order to compensate for level changes due to the IRN generation and the amplitude modulation, all digital stimuli where equalized to have the same rms level. The presentation level was then randomized by ±6 dB using a TDT PA4 programmable attenuator, in

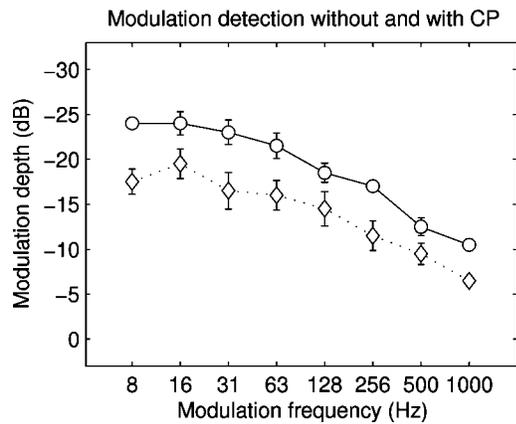Stein *et al.*: Interaction between carrier and envelope frequency

FIG. 2. The effect of carrier periodicity on amplitude-modulation detection: Circles and the solid line show modulation-detection thresholds with a Gaussian-noise carrier as a function of the modulation frequency. Diamonds and the dotted line show modulation-detection thresholds with an IRN$(d,1,16)$ carrier where the IRN delay, d, was equal to the reciprocal of the modulation frequency. Throughout the range of the tested modulation frequencies, the IRN carrier resulted in an impairment of modulation detection by about 5 dB.

order to eliminate residual level cues. The listeners were seated individually in a sound attenuating booth.

b. Procedure. Modulation detection thresholds were obtained using an adaptive two-interval, two-alternative, forced-choice procedure. The listeners' task was to detect the amplitude modulation of one randomly selected interval. Visual feedback was provided after every decision. The adaptive parameter was the modulation index, m, which was varied in terms of dB $(20 \log m)$. Its value was changed following a three-down, one-up rule. At the beginning of a run, the modulation index was changed in steps of 5 dB. After the third reversal, the step size was reduced to 2 dB, and after the fifth reversal, it was set to 1 dB. The mean over the following six reversals was taken as the threshold value for that run. This procedure estimates the 79.4% correct point of the psychometric function (Levitt, 1971). For each threshold run, one modulation frequency was randomly selected out of the whole range. Subjects were trained on the stimuli until they reached stable performance, i.e., until the standard error for the performance of three consecutive runs per signal configuration was no more than 2 dB. The mean of these last three runs per signal configuration was taken as the final threshold estimate.

c. Listeners. Four listeners (two female, two male) ranging in age from 25 to 35 years took part in Experiment I. One of them was the last author (LW). All subjects participated voluntarily and had clinically normal hearing.

3. Results

Figure 2 shows the threshold modulation depth as a function of the modulation frequency. The plot represents the average performance of all four listeners. The error bars indicate the standard error of the mean across listeners. The circles indicate data for the GN carrier, i.e., detection thresholds for AM in the absence of CP. The diamonds indicate data for the IRN carrier, i.e., detection thresholds for AM in the presence of CP.

For the GN carrier the subjects' ability to detect amplitude modulation decreases continuously as the modulation frequency increases. This finding is in line with the results obtained in previous studies (Eddins, 1993; Eddins, 1999; Ewert and Dau, 2000; Viemeister, 1979). For the IRN carrier condition, the overall shape of the threshold curve is the same, but it is shifted downwards by, on average, 5 dB. These results indicate that the sensitivity to AM is impaired by the presence of CP independent of the modulation frequency.

B. Experiment II: The effect of AM on CP detection

1. Rationale

Experiment II serves as a counter experiment to Experiment I. Listeners were asked to detect CP in a broadband noise carrier. Since the IRN gain factor allows changing the degree of CP continuously, this feature enables quantitative measurements of perceptual CP thresholds. To investigate the effect of AM on CP detection, the experiment was performed in the absence and in the presence of a sinusoidal AM imposed in both presentation intervals of a trial.

2. Method

a. Stimuli. The task was to discriminate an IRN stimulus from a GN stimulus. The IRN stimuli were generated with f0's between 7.8125 and 1000 Hz (equivalent to IRN delays between 128 and 1 ms) in octave steps. In order to change the degree of CP, the IRN gain was changed adaptively. The number of IRN iterations was fixed at 16. The imposed AM had the same frequency as the carrier f0 and a modulation index of 0 dB (100%). The presentation level and apparatus were the same as in Experiment I.

b. Procedure. The experimental procedure was equivalent to Experiment I, only differences are reported here. The adaptive parameter was the gain factor of the IRN. The step sizes were 10 dB for the first two reversals, 5 dB for reversals three to five, and 2 dB after the fifth reversal. For each threshold run, one f0 was randomly selected out of the whole range.

c. Listeners. Four listeners (two female, two male) ranging in age from 28 to 35 years participated in Experiment II. Except for listener LW, the listeners were different from those in Experiment I and included the first author (AS). All listeners had clinically normal hearing and participated voluntarily in this study.

3. Results

Figure 3 shows the threshold IRN gain required to detect carrier periodicity as a function of the IRN f0. The plot represents the average performance of all four listeners and the error bars indicate the standard error across listeners. The circles and solid line represent CP detection in the absence of AM; the diamonds and dotted line represent CP detection in the presence of AM.

In the absence of AM, the ability to detect CP improves with increasing f0 and reaches a plateau for an f0 of 125 Hz and above. In the presence of AM, sensitivity for carrier periodicity appears not to be affected by the AM for an f0 of

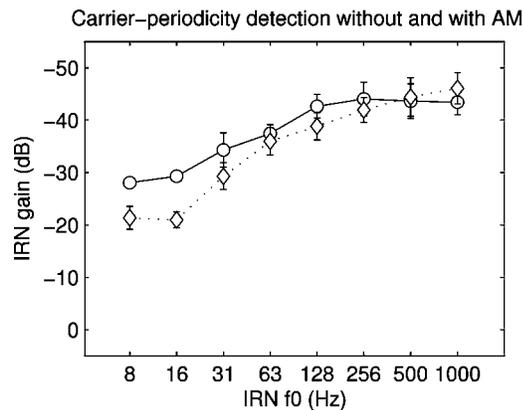**Carrier–periodicity detection without and with AM**

FIG. 3. The effect of amplitude modulation on carrier-periodicity detection: Circles and the solid line show carrier-periodicity detection thresholds without additional amplitude modulation as a function of the IRN f0. Diamonds and the dotted line show carrier-periodicity detection thresholds with a 0 dB SAM imposed on both the signal and standard stimuli. Additional amplitude modulation impaired sensitivity to carrier periodicity only for f0's below the pitch range (7.8125 and 15.625 Hz).

31.25 Hz and above. Only for the lowest f0's of 15.625 and 7.8125 Hz, the sensitivity to CP is impaired by the presence of AM.

III. DISCUSSION

The experimental results show that sensitivity to AM is significantly impaired by the presence of CP. However, sensitivity to CP is not impaired by the presence of AM as long as the carrier f0 is 31.25 Hz or above. As pointed out in Sec. I, CP as well as AM can evoke a pitch perception and it is conceivable that pitch and pitch strength are important perceptual cues used in these experiments. However, pitch is only perceived when the f0 is higher than about 30 Hz (Krumbholz *et al.*, 2000; Pressnitzer *et al.*, 2001).

In the following, the experimental results are discussed with respect to the possible role of pitch perception. The discussion is focused on two regions, the infra-pitch region, represented by the stimulus f0 of 7.8125 and 15.625 Hz, and the pitch region, represented by f0's of 31.25 Hz and above.

The lower limit of pitch coincides with the deviation of the performances for CP detection in Experiment II. This means that in the range of f0's eliciting a pitch sensation, sensitivity to CP is not impaired by additional AM. In the infra-pitch region, the thresholds for CP detection are significantly reduced. Experiment II shows that the IRN carrier periodicity remains well detectable for f0's of 7.8125 and 15.625 Hz (IRN delays of 128 and 64 ms). These low f0's result in the perception of a repeating feature rather than pitch. It appears that the listeners were confused whether this repeating feature arose from CP or AM. This assumption would result in impaired CP detection in the presence of AM as it was observed in Experiment II.

A confusion of CP and AM may also underlie the impaired AM detection in the presence of CP for the low f0's of 7.8125 and 15.625 Hz in Experiment I. However, in Experiment I this impairment is also found throughout the whole pitch range, i.e., also for all f0's of 31.25 Hz and above. A possible explanation for the impairment of AM detection in

the pitch region may lie in the differences of pitch strength elicited by the carrier and modulator. Experiments on rate discrimination with IRN (Krumbholz *et al.*, 2000) and SAM noise (Formby, 1985) indicate that with 16 iterations, the pitch strength of IRN is much larger than the pitch strength of SAM noise. Assuming that listeners partly relied on pitch cues for AM detection in Experiment I, the detection of the weak SAM noise pitch was supposedly impaired in the presence of the relatively strong pitch of the IRN carrier. In contrast, the weak SAM noise pitch was unable to impair the detection of the relatively strong IRN pitch in Experiment II for f0's of 31.25 Hz and above.

IV. MODEL SIMULATIONS

In the following, model simulations of the data in previous experiments are presented. The focus lies on the comparison of different implementations of periodicity extraction. Two model approaches are compared that share a common preprocessing and decision stage. The first model is a modulation filterbank (MFB) based on the model of effective signal processing by Dau *et al.* (1997a, b). The second, an autocorrelation (AC) model, is based on a model suggested by Meddis and O'Mard (1997). The most important difference between both model structures lies in the functional implementation of the way the auditory system extracts the periodicity (see Sec. I). Except for the different forms of periodicity extraction, both models have slight differences as the adaptation loops in the MFB model and outer- and middle-ear filtering as well as a detailed inner hair cell stage in case of the AC model. For the current study, the existing models were rearranged in a way that allows for a direct comparison between the two different forms of periodicity extraction.

A. Model structure

1. Peripheral processing

To account for effects of auditory filtering, the signals were subjected to a 20-channel gammatone filterbank with center frequencies equally spaced on a logarithmic axis between 100 and 6000 Hz. Effects of outer- and middle-ear filtering were not considered. After passing the gammatone filterbank, the internal representation of the stimuli was half-wave rectified, compressed with an exponent of 0.4 and subsequently low-pass filtered with a 1000 Hz, second-order Butterworth filter. These stages simulated the mechano-neural transduction of the inner hair cells. At this level, where the model's internal representation can be interpreted as representing the spike probability functions in the auditory nerve, either a modulation filterbank or an autocorrelator was applied.

2. Periodicity extraction

In one case, a modulation filterbank was included for each auditory frequency channel. The modulation filters were first-order resonant filters with a Q value of 2 as described in Dau *et al.* (1997a, b). Modulation filters were equally spaced on a logarithmic envelope-frequency axis, overlapping at their -3 dB points. They ranged from 10 Hz to a quarter of

the auditory-filter center frequency. The upper restriction of the center frequency is necessary, since the modulation spectrum which results from the process of half-wave rectification, compression and low-pass filtering, still contains information about the audio-frequency spectrum. The restriction ensures that only envelope information at the output of the respective auditory filters is extracted from the stimulus. For auditory filter center frequencies greater than 4000 Hz, the maximum modulation-filter center frequency was restricted to 1000 Hz. For modulation frequencies below 10 Hz, an additional second-order Butterworth modulation low-pass filter with a cut-off frequency of 7.5 Hz was applied. Modulation filtering was implemented in the modulation frequency domain by multiplying the envelope magnitude spectrum of a specific (audio-) frequency channel with the magnitude transfer function of the desired modulation filter. Only the rms value of the filter's envelope was considered. Thus, the time dimension of the model is lost and the model output is two-dimensional along the dimensions of audio frequency and modulation frequency. To restrict the model sensitivity to the amount of AM in the stimulus, by factoring out sensitivity to overall sound-level changes (see Sec. IV B 5), each bandpass modulation filter output was normalized to the corresponding low-pass modulation filter output which is equivalent to the average degree of excitation (level) in the respective auditory channel. Subsequently, the outputs of the modulation low-pass filters, which represent a spectral excitation pattern across audio-frequency channels, were discarded.

The AC model was implemented as a long-term autocorrelation. This is justified since the stimuli were time invariant and the time constant of a running autocorrelation is only a critical parameter for time variant stimuli (Wiegrebe, 2001). Thus the simulation output is two dimensional as in the MFB model, in this case along the dimensions of audio frequency and correlation lag. The AC model makes use of level differences in the spectral excitation pattern. To allow for the simulation with a roving-level paradigm, the model had to be made level independent (see Sec. IV B 5). Therefore the autocorrelation in each auditory channel was normalized, i.e., correlation coefficients were calculated. The resulting autocorrelation function was multiplied with a linear weighting function as suggested by Pressnitzer et al. (2001). The maximum correlation lag was 33 ms. In summary, the AC model used in the present study deviates from the model by Meddis and O'Mard (1997) in the following aspects: (a) The Meddis hair cell is replaced by a module which contains half-wave rectification, compression, and low-pass filtering, (b) the autocorrelation is normalized, and (c) the autocorrelation is not calculated in a short time window at the stimulus end. Instead a long-term autocorrelation over the whole stimulus duration is used.

3. A spectral model

Purely spectral models operate on the evaluation of the degree of excitation in each auditory frequency channel. Effectively, the rms at the output of the modulation low-pass filter from the MFB model is the average power in each auditory channel. By using only the modulation low-pass

filter, the output of the model corresponds to an auditory spectral excitation pattern. Therefore this model representation is comparable to purely spectral "pattern matching" pitch models like, e.g., Cohen et al. (1995), Goldstein (1973), and Terhardt (1974).

4. The decision device

The decision device was implemented as an optimal detector (Dau et al., 1996). Model decisions were based on the cross correlation between a template derived from 20 suprathreshold signal presentations of the two presentation intervals. The overall sensitivity can be adjusted by the variance of internal noise [see Dau et al. (1996) for details]. For the simulation using the AC model, the variance of internal noise was adjusted to achieve a predicted gain threshold of −45 dB for carrier periodicity detection of IRN with an f0 of 250 Hz. The internal noise variance was kept constant for all simulations using the autocorrelation stage. For simulations using the modulation filterbank, the variance of internal noise was set to a value for which the model predicted a threshold modulation depth of −16 dB for a modulation frequency of 15.625 Hz applied to a GN carrier. Within the MFB model, the AM detection in the region of −16 dB is limited by the external statistics of the broadband Gaussian noise carrier. The internal noise was set to a small value, with negligible influence on AM thresholds. Again, this value was the same for all simulations using the modulation filterbank. Model predictions were obtained using the same adaptive two-alternative, forced-choice procedure as in the experiments.

B. Simulation results and discussion

The simulation results of the modulation-filterbank model and the AC model are shown together with the experimental data in the top and bottom row of Fig. 4, respectively. The bold lines indicate the empirical data. The basis for the model performances is discussed with respect to the templates generated by the decision device. Some of these templates are shown in Figs. 5 and 6.

1. CP detection in the MFB model

The MFB model provides reasonable predictions of CP detection for intermediate f0's around 125 Hz [open circles in panel (a) of Fig. 4]. The templates for these f0's [e.g., Fig. 5(a)] indicate that the MFB model relies mostly on the detection of a modulation magnitude peak (dark areas) equal to f0 and an area of reduced modulation magnitude (dip, indicated by the brighter areas) for modulation filters tuned below the f0 of the stimulus [see the arrow in Fig. 5(a)]. The rippled structure of the IRN envelope spectrum is only partially resolved below the first harmonic by the modulation filterbank. The MFB model cannot correctly predict CP detection for the lowest f0 of 8 Hz. This deficit results from the dc sensitivity of the modulation filters. The dc response of the filters essentially swamps the dip-peak structure in the envelope magnitude spectrum up to f0. When additional AM was applied to the IRN carrier [open diamonds in panel (a) of Fig. 4], the MFB model performance is unaffected by the addition of the AM to the IRN for all f0's greater than 16 Hz.
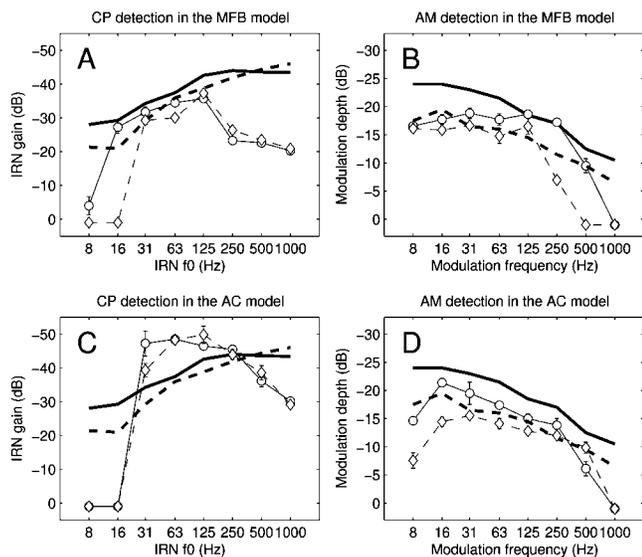
FIG. 4. Simulation of CP and AM detection in the modulation-filterbank (MFB) model (upper row) and in the autocorrelation (AC) model (lower row). The experimental data are replotted from Figs. 2 and 3 with bold lines. The solid lines represent data for CP and AM detection without interfering AM or CP. The dashed lines represent data for CP and AM detection with interfering AM or CP. The symbols connected by the fine lines show the simulation results. Open circles represent simulated data for CP detection without AM or AM detection without CP. Open diamonds represent simulated data for CP detection in the presence of AM or AM detection in the presence of CP. The error bars represent standard errors across three simulations.

For 8 and 16 Hz the performance breaks down completely. The templates for these two conditions are shown in Figs. 5(c) and (e). Note that the dip region in the modulation magnitude pattern [white area in Fig. 5(c)] is absent with additional AM, as seen in Fig. 5(e). This confirms the idea that the MFB model relies on the detection of the first dip in the envelope spectrum for CP detection.

The MFB model considerably underestimates CP detection performance for high $f0$'s above 125 Hz. In this case, the templates (not shown) reveal only spurious information on the reason for this lack in sensitivity. None of the activations in the templates correspond to either spectrally resolved harmonics or resolved peaks in the modulation spectrum. Note, in the present implementation, the MFB model cannot use audio-spectral information, since only the output of the modulation filters normalized by the overall power in the respective audio-frequency channel was considered. The output of the modulation low-pass was omitted.

2. AM detection in the MFB model

The MFB model predicts AM sensitivity quite well for the broadband noise carrier [open circles in Fig. 4(b)]. The template for the AM detection for a modulation frequency of 125 Hz [Fig. 5(b)] shows a clear vertical ridge of activation in the modulation filter tuned to 125 Hz. This ridge is more prominent for higher auditory filter CFs because only these filters are broad enough to pass the modulation. While for low and intermediate modulation frequencies almost no effect of additional CP is predicted [open diamonds in Fig. 4(b)], a strong deterioration in AM detection is observed at a
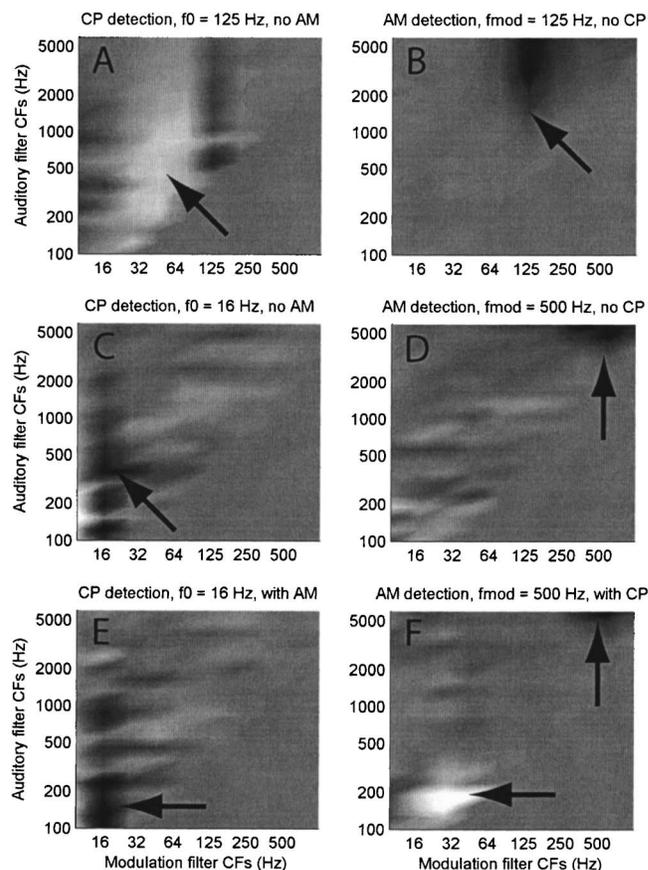

FIG. 5. Templates generated by the optimal detector connected to the output of the MFB model. Templates generated for CP detection (left column) and AM detection (right column) with various fundamental or modulation frequencies are shown. The arrows indicate prominent features in the templates that were used by the optimal detector to perform the CP or AM detection. See the text for further details.

modulation frequency of 500 Hz when the GN carrier is replaced by an IRN carrier. The corresponding templates are shown in Figs. 5(d) and (f). Figure 5(d) shows a small peak of excitation in the modulation filter tuned to 500 Hz which is only observed for the highest auditory-filter CFs. With an IRN carrier [Fig. 5(f)], the template reveals a pronounced dip in the modulation excitation pattern well below the modulation frequency and only for low auditory-filter CFs. This feature is highly reproducible and it appears to impair AM detection for these high modulation frequencies. For both carrier types no sensitivity to an AM of 1000 Hz is predicted. This is probably caused by the relatively narrow auditory filters at the highest center frequency of 6000 Hz as used in the model. They essentially do not pass AM for modulation frequencies above about 660 Hz. It is conceivable that the nonlinear human cochlear filters transmit more of these high-frequency modulations since they might be considerably broader at sound-pressure levels as used in the present study when compared to the linear gammatone filters used here.

3. CP detection in the AC model

Panel (c) of Fig. 4 shows thresholds for CP detection as predicted by the AC model, in absence and presence of additional AM (open circles and diamonds, respectively). Since

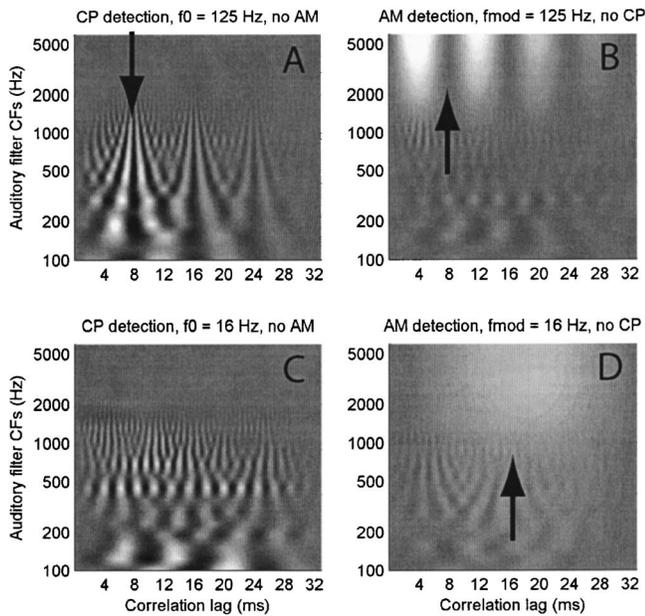Stein *et al.*: Interaction between carrier and envelope frequency

FIG. 6. Templates generated by the optimal detector connected to the output of the AC model. Comparable to Fig. 5, templates generated for CP detection (left column) and AM detection (right column) with various fundamental or modulation frequencies are shown. The arrows indicate prominent features in the templates that were used by the optimal detector to perform the CP or AM detection. See the text for further details.

the maximum autocorrelation lag was set to 33 ms, the AC model is unable to detect the CP of f0's lower than 30 Hz. When the f0 of the CP is 125 Hz [Fig. 6(a)], the CP causes a vertical ridge in the template at a correlation lag of 8 ms equal to the stimulus period [Fig. 6(a), arrow]. When the f0 is 16 Hz [Fig. 6(c)], this ridge falls outside the template, and the model is unable to detect the CP.

While the model shows better detection performance for intermediate carrier periodicities than observed in the data, the AC model is too insensitive for high f0's. Again, note that the AC model could not rely on spectral information since the autocorrelation was normalized in order to make the model level insensitive. In the correlation-lag dimension, the half-wave rectification and low-pass filtering impairs the encoding of the periodicity of the relatively high absolute frequencies of even the low harmonics.

The AC model correctly predicts virtually no influence of AM on CP detection.

4. AM detection in the AC model

Panel (d) of Fig. 4 shows good agreement between the predicted AM detection thresholds for the AC model and the empirical data. The good performance of the AC model in AM detection of low AM frequencies appears surprising since the maximum autocorrelation lag was limited. However, the autocorrelation of the model's internal representation of a sinusoidally amplitude modulated stimulus is also sinusoidal. The internal representation is similar to the envelope of the stimulus simulated by half-wave rectification and low-pass filtering in the inner hair cell stage. This is shown in the template of the AC model for the detection of an AM of 125 Hz [Fig. 6(b), arrow]. Thus AM detection is possible even when only a portion of the modulator period (e.g., the

first dip) fits into the autocorrelation function [Fig. 6(d)]. The current simulations show that this residual information in case of low AM frequencies provides a reliable cue for AM detection in the AC model.

A slight deterioration of AM sensitivity is observed in the presence of CP [open diamonds in Fig. 4(d)].

As in the MFB model, the degree of this deterioration appears to be too small when compared to the empirical data.

5. CP and AM detection in a spectral model

At first sight, it appears pointless to simulate AM detection on broadband carriers in a purely spectral model since the AM does not produce any changes in the long-term spectrum of a broadband stimulus. However, a spectral model provides very useful insights on the detection strategies of more complex auditory models as they are investigated here.

Simulation results for CP and AM detection in the spectral (modulation low-pass) model are shown in Fig. 7. As in the previous plots, the solid lines with open circles represent empirical data for CP and AM detection without interfering AM or CP. The dashed lines with open diamonds represent empirical data for CP and AM detection with interfering AM or CP. The top row shows the simulation results obtained with a roving level paradigm (±6 dB SPL), the bottom row shows the results for the same experiment without roving level. The simulation results in the top row show that CP detection is possible for higher IRN f0's where the audio-spectral peaks of the IRN become partially resolved [panel (a) of Fig. 7], as expected. For f0's of 125 Hz and below, this spectral ripple profile, as it is preserved after peripheral processing, appears not to be sufficient for detecting CP when a roving level paradigm is used. Without level roving [panel (c)] CP detection is well predicted for all f0's above 31 Hz.

Panel (b) of Fig. 7 shows that AM detection cannot be predicted at all when a level rove is applied to the signal. The simulation results in panel (d) of Fig. 7 reveal the importance of level roving. Without roving level, the spectral model becomes capable of AM detection. It is a well-described phenomenon that nonlinear peripheral auditory processing influences the degree of central auditory excitation (Gockel *et al.*, 2003; Kohlrausch and Sander, 1995). This problem also applies to the detection of AM in a spectral model: The AM sensitivity of the spectral model is based on the fact that the amplitude-modulated stimulus produces less auditory excitation than the unmodulated stimulus. The compressive nonlinearity has a stronger effect on the modulated stimulus than on the unmodulated stimulus. Note that the loss of excitation is a result of both the half-wave rectification and the power-law compression. The loss of excitation occurs although the rms value at the model input is the same for the modulated and for the unmodulated stimulus as in the present study. The effect of the compressive nonlinearity applies differently in auditory frequency channels with different center frequencies dependent on the amount of modulation that is passed by the filters. When the modulation frequency is very low (below about 64 Hz) all auditory filters are broad enough to pass the modulation and the activity is strongly reduced in all filters as a consequence of the compressive nonlinearity. For intermediate modulation frequencies, the strong effect of auditory
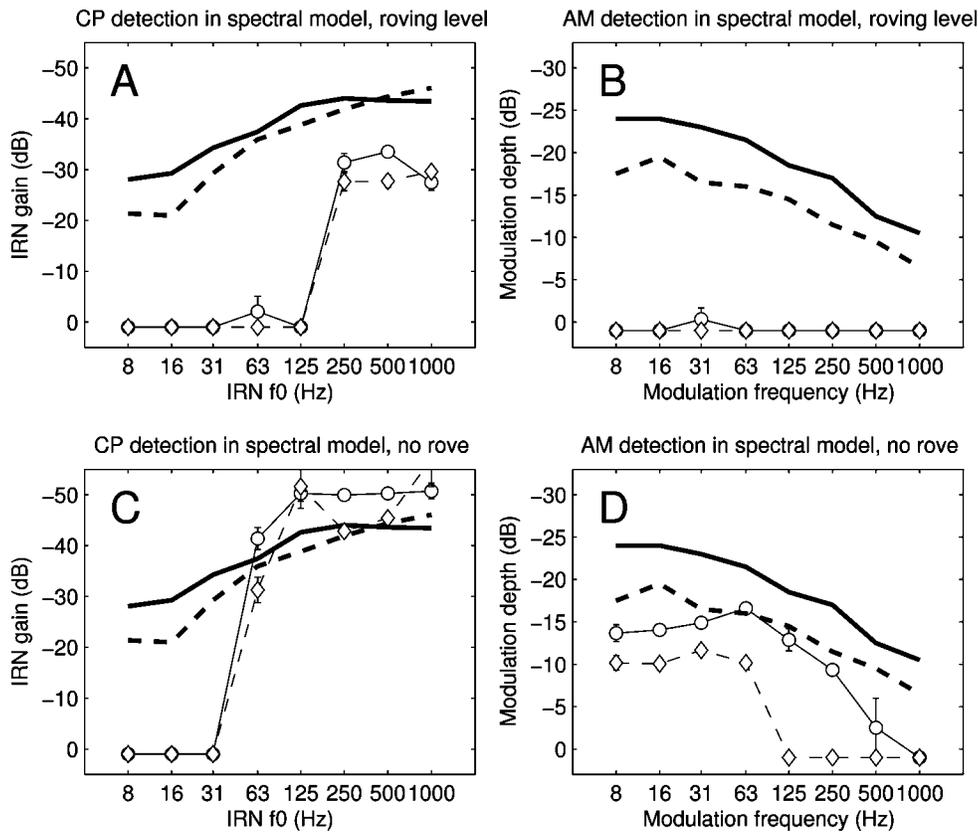
FIG. 7. CP detection (left column) and AM detection (right column) in a spectral model. The same line types and symbols as in Fig. 4 are used. In the upper row, simulation results were obtained with a roving-level paradigm, in the lower row, no level roving was used. The roving level had a dramatic effect on the model predictions. Without roving level, the spectral model was able predict AM sensitivity quite well although the long-term magnitude spectrum of the stimulus was not affected by the AM.

nonlinearity only applies to those filters with a center frequency higher than about eight times the modulation frequency. Very high modulation frequencies (1000 Hz) are transmitted by none of the auditory filters used for the current simulations. Thus, the broadband AM noise generates a characteristic auditory excitation pattern with reduced activity in channels that pass the modulation. This pattern serves as a spectral detection cue for the decision stage in AM detection experiments with broadband noise carriers. While the psychophysical relevance of these modeling results has not been formally investigated, it appears highly unlikely that listeners exploit such auditory cues, related to spectral profile analysis, to detect AM.

In order to study the effect of level roving with the MFB and the AC model, the internal representation of the models had to be level independent. Otherwise, the level roving would have completely deteriorated the models' performance as it would have acted as a strong source of "external" noise. Such an effect is clearly not observed in the empirical data. Effectively, the requirement of level independence removed the availability of purely spectral excitation cues in the MFB and AC model.

V. THEORETICAL RELATIONSHIP BETWEEN MODULATION FILTERS AND AUTOCORRELATION

The current paper assesses the relationship between a modulation filterbank analysis and an autocorrelation based analysis of the internal representation of auditory models at a stage comparable to the auditory nerve. Up to this point, modulation filterbank analysis and autocorrelation have been compared and treated as two independent modeling approaches to extract periodicity from the internal representation of the stimuli within the models. This section gives some insight into the mathematical relationship between the two processing schemes.

A modulation filter is described in the time domain by the convolution of the filter's impulse response and the internal representation, which is generally referred to as wave form in the following. A hypothetical impulse response (IR) of a modulation filter with a modulation center frequency of 62.5 Hz is given in the upper left panel of Fig. 8.

The autocorrelation function (ACF) of a wave form for a given time lag is mathematically described by summing the product of the wave form and a delayed version of it along the time axis. The ACF for a specific lag is often associated with a corresponding comb-filter type analysis of the stimulus spectrum, which, in this case, resembles a comb-filter type modulation filter. The impulse response of the respective comb filter is unipolar having two peaks separated by the lag ("U-shape," see upper right panel of Fig. 8). However, the convolution with such an impulse response is mathematically described by a summation of the stimulus and the time delayed version of it, clearly differing from the ACF, where the stimulus and its delayed version are first multiplied and then summed over time. A similarity between the two processes is introduced when the power or the variance of the comb-filter output is obtained and compared to the ACF for the respective lag. The computation of either quantity includes squaring of the filter output and successive summation. The squaring introduces a multiplicative cross-term to the filtering process which makes it comparable to the ACF. Figure 9 compares autocorrelation and convolution of

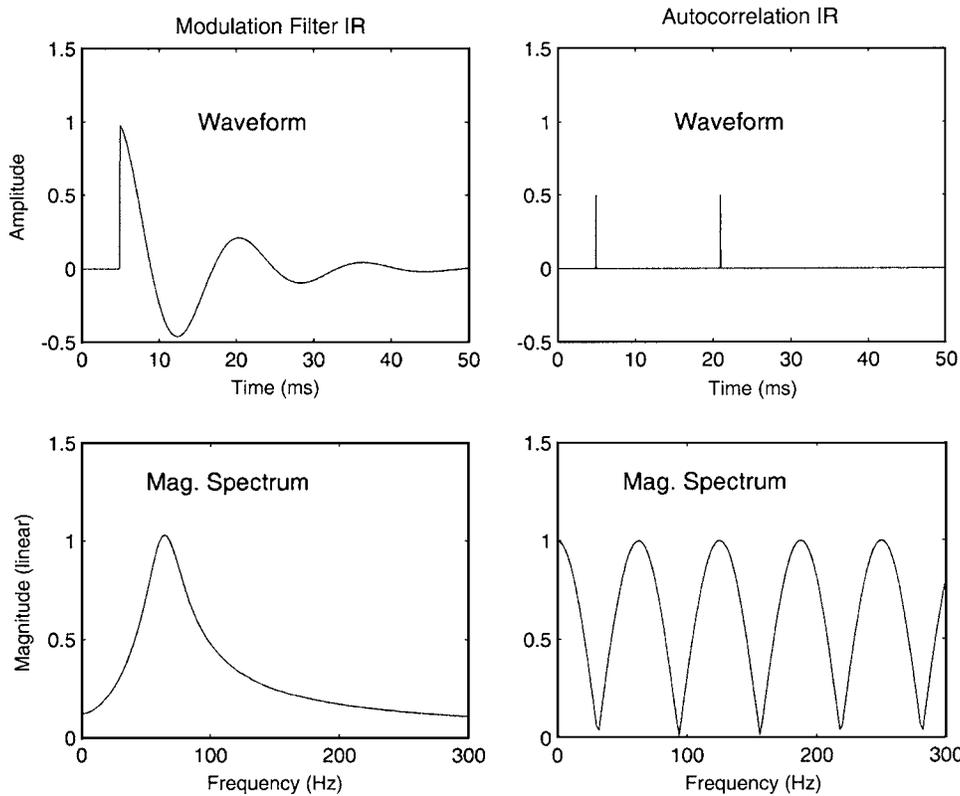Stein *et al.*: Interaction between carrier and envelope frequency

FIG. 8. Comparison of the impulse response of a modulation filter (left column) and the impulse response which produces results most similar to an autocorrelator (see Fig. 9). The respective transfer functions of the filters are shown in the lower row. While the modulation filter has a bandpass characteristic, the autocorrelation can be approximated with a "U-shaped" impulse response having a comb-filter transmission characteristic.
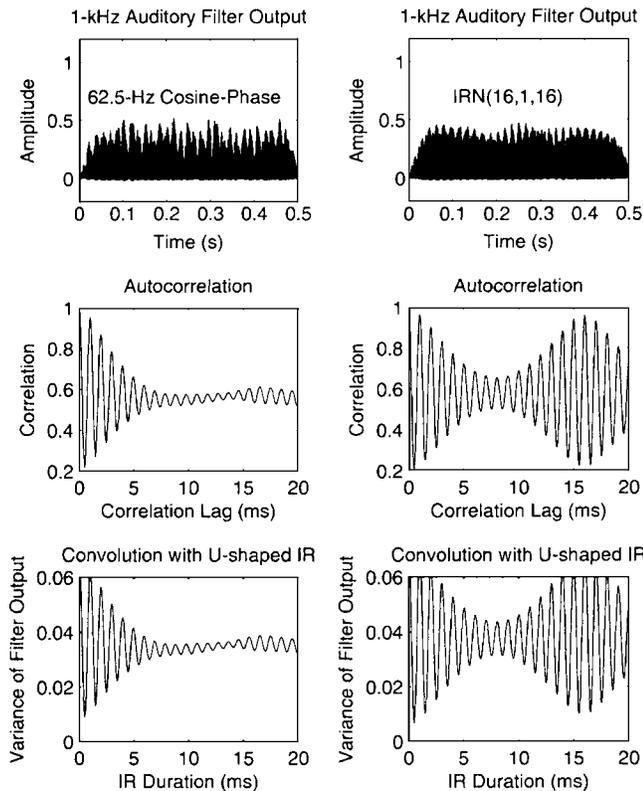

FIG. 9. Comparison of autocorrelation with the output of a filterbank of "U-shaped modulation filters." The input to the periodicity analysis consists of the wave form at the output of the 1 kHz auditory filter stage generated by either a 62.5 Hz SAM noise (left column) or by an IRN(16,1,16) with the same f0 (right column). The variance of the filterbank output (bottom row) leads to results that are very similar to the autocorrelation function (middle row).

the output wave form from a 1 kHz auditory frequency channel in response to either a 62.5 Hz SAM noise (left column) or in response to IRN(16,1,16). The auditory-filter, output wave forms (after half-wave rectification, compression, and low-pass filtering) are plotted in the upper row and the ACFs of these wave forms are shown in the middle row. The lower row shows the variance of a filterbank output produced with U-shape IRs having a duration equal to the correlation lag in the middle row. In fact, the variance of the U-shape modulation-filter outputs and the ACF are qualitatively similar. Thus, the difference between autocorrelation and a modulation-filterbank analysis can be reduced to three critical points.

First, with the observation of equally spaced correlation lags, the spacing of the fundamental frequencies of the comb filters is equal on a *linear* periodicity axis for an autocorrelator. For the modulation filterbank, the spacing is assumed to be equal on a *logarithmic* envelope-frequency axis.

Second, the IR representing the autocorrelation for a specific lag differs from that of a modulation filter as used in the present study. This difference is shown in Fig. 8. The magnitude spectra in the lower row reveal that unlike the bandpass modulation filter of the modulation filterbank (left), the autocorrelation IR has a rippled (comb-filter) spectrum (right). In the context of the purely temporal pitch extraction, this rippled spectrum is essential to capture the contribution of higher-order harmonics to the overall pitch. The modulation-filter magnitude-spectrum shows only a single peak.

Third, in the current implementation of the modulation filterbank, the highest modulation-filter center frequency was limited to a quarter of the center frequency of the preceding auditory filter. This was done to ensure that residual carrier

information is effectively removed from the internal representation of the envelope. In the templates shown in Fig. 5, this area where no modulation filters were applied can be seen as the uniformly grey triangle in the lower right of each template. With this restriction, the modulation filterbank model would be unable to extract the pitch of a sinusoid or the pitch of low harmonics in the time domain. In the autocorrelation approach, this restriction was ignored: The autocorrelation is calculated for a correlation lag of 10 ms for the output of an auditory filter centered at 100 Hz. Thus, it appears that the autocorrelation makes extensive use of information which has been deliberately excluded in the modulation-filter approach.

When modulation filters with center frequencies similar to the auditory-filter center frequencies were included in an additional test, the simulation results as shown in Fig. 4(a) improved for CP detection with high f0's of 250, 500, and 1000 Hz (not shown). AM detection was not affected because the AM is not transmitted by auditory filters tuned to relatively low frequencies. This approach, however, would add information to the MFB output which is not contained in the (Hilbert) envelope of the stimulus. Note that a 100 Hz pure tone sounds quite different from any percept generated by a high-frequency tone amplitude modulated at 100 Hz (apart from the quadratic distortion product). If there was a 100 Hz modulation filter analyzing the output of a 100 Hz auditory filter, this modulation filter would be strongly excited since the internal representation of the envelope in the model, unlike the Hilbert envelope, would contain pronounced carrier periodicity information. In the MFB model, this carrier periodicity could then not be separated from envelope periodicity.

VI. TOWARDS A UNIFIED MODEL OF MODULATION- AND PITCH PERCEPTION

It appears that the AC model and the modulation filter model share many properties that are essential for a successful extraction of carrier- and envelope periodicity. The remaining question is what the ingredients for a unified model should be. From the current study some manipulations to the MFB model which could turn it into a successful pitch model are suggested.

First, it appears that the dc sensitivity of the modulation filters as used in Dau *et al.* (1997a, b) should be omitted. This dc sensitivity can swamp the sensitivity to relatively smaller features in the envelope-frequency domain. The modulation filters suggested in Ewert and Dau (2000) and Ewert *et al.* (2002) show no dc sensitivity.

Second, modulation filters should be included with a modulation center frequency equal to the audio center frequency. These filters should have a harmonic ripple in their transfer function. This is mandatory to catch, in the envelope-frequency domain, the strong influence of higher-order spectrally resolved harmonics on pitch (e.g., the pitch equal to f0 caused by the third, fourth, and fifth harmonic). This rippled transfer function may only be necessary for modulation filters with modulation center frequencies above the lower limit of pitch. The transfer function of modulation filters as inferred from modulation masking experiments

does not contradict such a ripple fine structure, in particular if it is assumed that the ripples might be less pronounced than the first (center-frequency) peak of the modulation filter.

Regarding useful changes to the AC model, the linear periodicity axis as produced by the autocorrelation should be replaced by a logarithmic periodicity axis. Both perceptual evidence and physiological properties of the auditory brainstem render a linear periodicity axis highly unlikely.

It remains to be investigated how the modulation filterbank deals with inharmonically shifted stimuli: Related to the stimuli used in the current study, inharmonic stimuli can be generated by using a negative gain factor in the IRN delay-and-add loop. The pitch of IRN with negative gain differs substantially from the pitch produced with positive gain (Yost, 1997). This is the case although the envelope spectra of IRN with positive and negative gain are identical. Wiegrebe and Winter (2001a) have shown that, as a consequence, the spectrally unresolved harmonics of IRN are encoded in the same way independent of the sign of the IRN gain. As the current implementation of the modulation-filterbank model does not evaluate the spectral excitation pattern since the modulation low-pass was omitted, it would not be sensitive to the differences in the audio spectrum produced by the resolved harmonics of IRN with positive and negative gain. Thus, while a quantitative simulation of these effects is beyond the scope of this paper, it appears unlikely that the modulation-filterbank model in its present form can account for the perception of inharmonically shifted stimuli. However, the above-suggested improvements, especially the use of spectrally rippled modulation filters, may enable an improved model performance in this domain. Presumably, these rippled modulation filters will allow one to analyze the mismatch between carrier and envelope period which is typical for inharmonically shifted stimuli.

VII. SUMMARY AND CONCLUSIONS

The current experiments investigated interactions between AM perception and CP perception in broadband stimuli. The results showed that sensitivity to AM decreased significantly when the GN carrier was replaced by an IRN carrier whose f0 was equivalent to the modulation frequency. This was the case for all modulation frequencies tested.

In contrast, sensitivity to CP did not decrease when the stimuli were amplitude modulated with a modulation frequency equal to the f0, as long as the f0 was above the lower limit of pitch [about 30 Hz, Krumbholz *et al.*, (2000); Pressnitzer *et al.*, (2001)]. Below the lower limit of pitch, the presence of AM decreased sensitivity to CP. In this region, the interference might be related to a confusion effect of the repeating envelope pattern.

The simulation of the experimental results in a generic implementation of either the modulation filterbank or the autocorrelation model of periodicity extraction indicates that while both models capture the general trends of the data, none of the two provides a satisfactory explanation for AM and CP interaction. The direct comparison of the autocorrelation and the modulation-filter approach combined with a power or variance based decision variable reveals that the

basic difference lies in the transfer function of the modulation filter and the transfer function of the autocorrelator: Modulation filters have a bandpass transfer function while the filters which produce an output akin to autocorrelation have a comb-filter transfer function. The simulation of the experimental results with a spectral model shows that AM of broadband stimuli can produce changes in the spectral excitation pattern which, however, are very unlikely to underlie AM sensitivity.

From these simulations it appears promising to construct a hybrid temporal model with properties of both a modulation filter and an autocorrelator. It is intriguing that the temporal filtering properties of Chop-S units described in the ventral cochlear nucleus of many mammals can be regarded as such a hybrid of a modulation filter and an autocorrelator (Wiegrebe and Winter, 2001b; Winter et al., 2001). Thus, the parallel pursuit of both physiological and functional models of periodicity encoding (like the present ones) may lead to a unified understanding of the auditory extraction of periodicity and the perception of pitch.

ACKNOWLEDGMENTS

Burns, E. M., and Viemeister, N. F. (**1976**). "Nonspectral pitch," J. Acoust. Soc. Am. **60**, 863–868.

Burns, E. M., and Viemeister, N. F. (**1981**). "Played-again Sam: Further observation on the pitch of amplitude- modulated noise," J. Acoust. Soc. Am. **70**, 1655–1660.

Carlyon, R. P. (**1998**). "Comments on 'A unitary model of pitch perception' [J. Acoust. Soc. Am. **102**, 1811–1820 (**1997**)]," J. Acoust. Soc. Am. **104**, 1118–1121.

Cohen, M. A., Grossberg, S., and Wyse, L. L. (**1995**). "A spectral network model of pitch perception," J. Acoust. Soc. Am. **98**, 862–879.

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., Puschel, D., and Kohlrausch, A. (**1996**). "A quantitative model of the 'effective' signal processing in the auditory system. I. Model structure," J. Acoust. Soc. Am. **99**, 3615–3622.

Eddins, D. A. (**1993**). "Amplitude modulation detection of narrow-band noise: Effects of absolute bandwidth and frequency region," J. Acoust. Soc. Am. **93**, 470–479.

Eddins, D. A. (**1999**). "Amplitude-modulation detection at low- and high-audio frequencies," J. Acoust. Soc. Am. **105**, 829–837.

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

Ewert, S. D., Verhey, J. L., and Dau, T. (**2002**). "Spectro-temporal processing in the envelope-frequency domain," J. Acoust. Soc. Am. **112**, 2921–2931.

Formby, C. (**1985**). "Differential sensitivity to tonal frequency and to the rate of amplitude modulation of broadband noise by normally hearing listeners," J. Acoust. Soc. Am. **78**, 70–77.

Gockel, H., Moore, B. C., and Patterson, R. D. (**2003**). "Asymmetry of masking between complex tones and noise: Partial loudness," J. Acoust. Soc. Am. **114**, 349–360.

Goldstein, J. L. (**1973**). "An optimum processor theory for the central formation of the pitch of complex tones," J. Acoust. Soc. Am. **54**, 1496–1516.

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

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.

Krumbholz, K., Patterson, R. D., and Pressnitzer, D. (**2000**). "The lower limit of pitch as determined by rate discrimination," J. Acoust. Soc. Am. **108**, 1170–1180.

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

Licklider, J. C. R. (**1951**). "A duplex theory of pitch perception," Experientia **7**, 128–133.

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

Meddis, R., and Hewitt, M. J. (**1991b**). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery.II. Phase sensitivity," J. Acoust. Soc. Am. **89**, 2883–2894.

Meddis, R., and O'Mard, L. (**1997**). "A unitary model of pitch perception," J. Acoust. Soc. Am. **102**, 1811–1820.

Pressnitzer, D., Patterson, R. D., and Krumbholz, K. (**2001**). "The lower limit of melodic pitch," J. Acoust. Soc. Am. **109**, 2074–2084.

Terhardt, E. (**1974**). "Pitch, consonance, and harmony," J. Acoust. Soc. Am. **55**, 1061–1069.

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

Wiegrebe, L. (**2001**). "Searching for the time constant of neural pitch extraction," J. Acoust. Soc. Am. **109**, 1082–1091.

Wiegrebe, L., and Patterson, R. D. (**1999**). "The role of envelope modulation in spectrally unresolved iterated rippled noise," Hear. Res. **132**, 94–108.

Wiegrebe, L., and Winter, I. M. (**2001a**). "Psychophysics and physiology of regular-interval noise: Critical experiments for current pitch models and evidence for a 1st-order, temporal pitch code," in *Physiological and Psychophysical Bases of Auditory Function*, edited by J. Breebaart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastricht).

Wiegrebe, L., and Winter, I. M. (**2001b**). "Temporal representation of iterated rippled noise as a function of delay and sound level in the ventral cochlear nucleus," J. Neurophysiol. **85**, 1206–1219.

Winter, I. M., Wiegrebe, L., and Patterson, R. D. (**2001**). "The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig," J. Physiol. (London) **537**, 553–566.

Yost, W. A. (**1996**). "Pitch of iterated rippled noise," J. Acoust. Soc. Am. **100**, 511–518.

Yost, W. A. (**1997**). "Pitch strength of iterated rippled noise when the pitch is ambiguous," J. Acoust. Soc. Am. **101**, 1644–1648.

Yost, W. A., Patterson, R., and Sheft, S. (**1998**). "The role of the envelope in processing iterated rippled noise," J. Acoust. Soc. Am. **104**, 2349–2361.