In: Hearing- Physiological Bases and Psychophysics, R. Klinke and R. Hartmann Eds, Springer, Berlin, 1983

MONOTIC AND DICHOTIC PITCH JND'S COMPARED

F.A. Bilsen and J. Raatgever

Applied Physics Department,
Delft University of Technology, the Netherlands

1. INTRODUCTION

In the peripheral auditory system, complex sounds are decomposed into a series of sinusoidal frequency components. These components can give rise to individual pitches, or to a low pitch if a more or less harmonic relation exists between them. Information on these components, either spectral (place) or temporal, is carried on to the central nervous system and constitutes the "central spectrum". Modern pitch theories (Goldstein, 1973; Terhardt, 1972; Wightman, 1973) have in common that the central pitch processor performs an operation on this central spectrum.

Reviewing the pitches of monotic and dichotic noise signals, Bilsen (1977) indicated that parsimony of pitch processing calls for utmost three ways in which a central spectrum is generated: (a) directly from the spectral information at either of the cochleae (internal, peripheral spectrum), (b) directly from the spectral information of both cochleae together (Houtsma and Goldstein, 1972), and (c) after binaural processing on the temporal information from left and right cochlea (Bilsen and Goldstein, 1974). Here, we are faced with important questions about the relative rôle of temporal versus spectral coding, and about the limitations on processing set by the auditory system.

Pertinent to these issues is the accuracy of pitch perception as expressed in the just noticeable difference of pitch (JND). Therefore, in the present paper, the JND is compared for different types of signals, viz. monotic signals like the periodic pulse (PP) and comb filtered noise (COMB) and dichotic signals like multiple-phase-shifted noise (MPS). It will appear possible to predict the MPS-JND from the COMB-JND by considering the shape of the internal spectrum as measured psychophysically, its internal modulation depth M_{int} in particular.

2. HYPOTHESES

In order to compare the JND's of monotic and dichotic signals and to predict one from the other, hypotheses, some of which have been tested already, have to be made, viz.:
- Signals with similar (internal) power spectrum evoke the same pitch (much evidence available; compare e.g. Bilsen, 1977).
- Signals with similarly shaped (internal) power spectrum having the same peak-valley difference, i.e. modulation depth M_{int} (dB), give the same JND.
- The internal spectrum of dichotic noise signals can be found from appropriate BMLD-measurements (compare Raatgever and Bilsen, 1977).
- The M_{int} (dB) of COMB-signals is equal to $M_{acoust.}$ (dB) for the first harmonics (compare Houtgast, 1977).

3. SIGNAL DESCRIPTION

The use of a periodic pulse (PP) as a test stimulus in the JND experiments being sufficiently familiar, COMB-noise and MPS-noise require further explanation.

COMB-noise is a monaural stimulus. It is the result of feeding white gaussian noise into a delay line with feed-back of a fraction g of the output. This results in a comb-like power spectrum with peaks at harmonic frequencies n/τ as indicated in Fig. 1b, where τ represents the delay (n=0,1,2...). The figure shows normalized power spectra of COMB-noise for two values of the feed-back factor g. It clearly demonstrates the influence of g on the modulation depth, characteristic for COMB-noise. Theoretically, the peaks in the power spectrum have the value $1/(1-g)^2$, whereas the minimal values in between are given by $1/(1+g)^2$.

MPS-noise is a dichotic stimulus in the sense that no spectral information is present at the separate ears but the desired spectral patterns arise due to binaural interaction. Although detailed knowledge is necessary to predict the internal spectra from their dichotic signal conditions, we use some essential simplifications to illustrate the spectral patterns involved (Raatgever, 1980). We consider the perceptual image of such dichotic noise stimuli, lateralized in the centre of the head, to be the result of a perfect addition of the signals at both ears, under preservation of their time structures. The, in this way idealized, internal signals with their specific spectral properties are assumed to constitute the spectral inputs for the pitch extracting mechanism in an analogous way as assumed for monaural signals like PP and COMB-noise. In the dichotic stimuli, moreover, the spectral information is subject to imperfections of the binaural interaction mechanism.

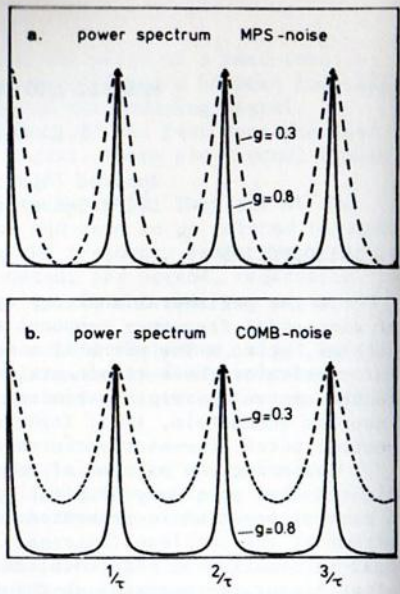

Fig. 1. Normalized power spectra (linear scale) of MPS-noise (a) and COMB-noise (b) for g=0.3 and 0.8

MPS-noise is, like COMB-noise, the result of white gaussian noise passing a delay line with feed-back of a signal fraction g. In order to obtain a flat frequency spectrum a negative fraction of the input $-g/(1-g^2)$ is added to the output (Bilsen, 1976). The signal at the contralateral ear is the original noise, balanced for equal power at both ears. Due to the specific phase relation between the signals at the ears, the resulting internal power spectrum of MPS-noise has peaks at harmonic frequencies n/τ as shown in Fig. 1a ($n=0,1,2...$). In this figure the supposed normalized power spectrum of MPS-noise is shown for two values of g. Theoretically, the peaks are given by $4/(1-g)^2\ (1+g)^2$. For frequencies just in between, the power equals zero for every g. This difference between the internal power spectra of COMB-noise and MPS-noise is reduced by the imperfect, noisy, operation of the binaural system, effectively reducing the modulation depth.

4. EXPERIMENTAL PROCEDURE

Two sets of JND-experiments have been carried out. These experiments concerned the monotic COMB-noise at one hand and the dichotic MPS-noise at the other. Both sets of experiments have been performed using four normal hearing subjects. Two of the subjects took part in all experiments.

In all cases a two-alternative forced-choice procedure has been used to measure the JND of pitch. The JND thus obtained corresponds to a 75% correct discrimination in f_o. For COMB-noise as well as for MPS-noise the JND is defined as: $JND=\Delta f_o/f_o$ (%). Here, $f_o(=1/\tau)$ is the frequency of the fundamental in the spectrum and τ is the basic delay used in the generation of the signals (section 3). JND-measurements were based on the detection of τ. For JND-measurements, in general, the condition $\Delta\tau<<\tau$ is valid, so in that case: $JND\approx \Delta\tau/\tau$ (%).

Subjects, seated in a sound-proof booth, were presented with stimuli with an average sensation level of 40 dB. TDH 39 headphones have been used. The signals always consisted of two properly shaped stimuli of 500 ms duration.

The subjects had to respond to a number of series of maximally 12 pairs of randomized stimuli with fixed $\Delta\tau$. Each series resulted in a positive score if at least 75% of the responses were correct. During the monotic experiments these scores have been recorded systematically as a function of $\Delta\tau$, i.e. the spectral difference to be detected. The JND's have been estimated from the average reversal from positive to negative scores. For the dichotic experiments a computer-controlled adaptive method has been used to determine the JND of dichotic pitch from the positive and negative scores in the same way. In both experiments the 90%-confidence limits have been computed.

5. RESULTS

a) *JND of COMB-noise*

The JND of COMB-noise was measured as a function of f_o $(=1/\tau)$, for three values of g, by four subjects. The results of two subjects which also participated in the dichotic JND-measurements, are represented in Fig. 2. The bars in the figures indicate the 90%-confidence limits. The other two subjects showed similar results. For f_o=200 Hz (τ=5 ms), the JND is replotted as a function of g for these four subjects in Fig. 4.

Note that for g=0.85 the JND is independent of f_o and that the average value for the four subjects amounts to 0.3%. For smaller values of g, the JND increases and becomes dependent on f_o.

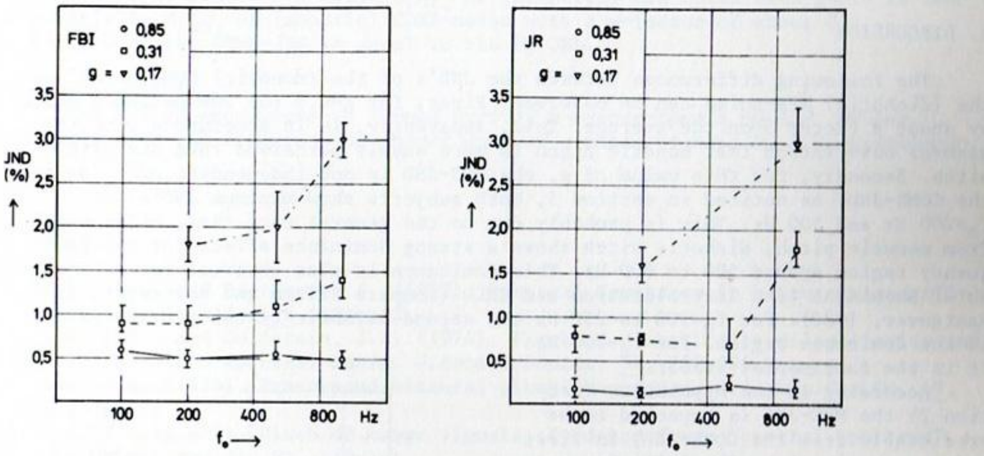

Fig. 2. The just noticeable difference in pitch (JND) of COMB-noise as a function of $f_o(=1/\tau)$ for three values of g and two subjects

b) *JND of MPS-noise*

The results of the JND-experiments for MPS-noise are plotted in Fig. 3 for g=0.8 and two subjects. The figures show the relative JND as a function of $f_o(=1/\tau)$. The 90%-confidence limits are indicated in the figures.

For both subjects we see minimal JND's for frequencies of the fundamental around 200 to 250 Hz and around 500 Hz. The minimal JND is 0.75% for observer JR and 1% for FB. Maximal JND's of about 2.3% are found for both observers at frequencies around 350 to 400 Hz. These two observers took also part in the monaural experiments. The other two subjects, not presented here, showed a similar behaviour, although one had significantly higher JND's.

The shape of the JND-curves presented here is different from the monaural data. It seems that JND's of dichotic pitch are minimal if the fundamental or its second harmonic is in the range of 500 Hz.

JND-measurements have also been performed with an extra interaural delay resulting in a lateralized pitch image of MPS-noise. The results, not presented here, show strikingly little effect on the JND's.

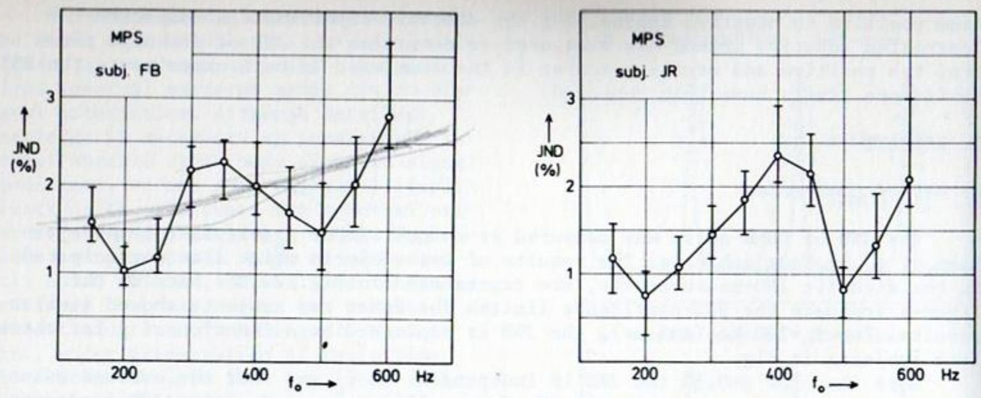

Fig. 3. The just noticeable difference in pitch (JND) of MPS-noise as a function of f_o (=1/τ) for g=0.8 and two subjects

6. DISCUSSION

The following differences between the JND's of the (monotic) COMB-noise and the (dichotic) MPS-noise can be observed. First, for g=0.8 the 200 Hz-JND's differ by about a factor 3 on the average. This, apparently, is in accordance with the general observation that monotic pitch is more easily perceived than dichotic pitch. Secondly, for this value of g, the MPS-JND is not independent of f_o as is the COMB-JND. As noticed in section 5, both subjects show minimum JND's for f_o=200 Hz and 500 Hz. This is probably due to the general fact that, different from monotic pitch, dichotic pitch shows a strong dominance effect for the frequency region around 500 to 600 Hz. This dominance is also observed for other binaural phenomena like lateralization and BMLD (compare Bilsen and Raatgever, 1973; Raatgever, 1980). For f_o=200 to 250 Hz the second harmonic of the MPS-spectrum is in the dominance region. For f_o=500 Hz it is the fundamental itself.

According to the hypotheses (section 2) the MPS-JND is expected to be predictable from the COMB-JND, in first order approximation. The following reasoning applies: For the MPS-signal the internal spectrum is reflected by appropriate BMLD-measurements. Thus, the modulation $M_{int.}$(dB) of the internal spectrum can be taken as the average peak-valley difference in dB for the (most pronounced) first and second peak. For f_o=250 Hz and g=0.8, this value amounts to 7 dB for subject JR and 8 dB for subject FB (Raatgever and Bilsen, 1977, Fig. 2). For COMB-noise $M_{int.}$(dB)= $M_{acoust.}$(dB) for the lowest peaks (compare Houtgast, 1977). Using this equality we are able to find the (monotic) COMB-spectrum that shows the same modulation depth as the MPS-signal. With the relation $M_{acoust.}=10\ \log\{(1+g)^2/(1-g)^2\}$ (compare section 3) it follows that g=0.38 for subject JR and g=0.43 for subject FB. Now, in Fig. 4 we read the corresponding JND-values, viz. 0.7% for

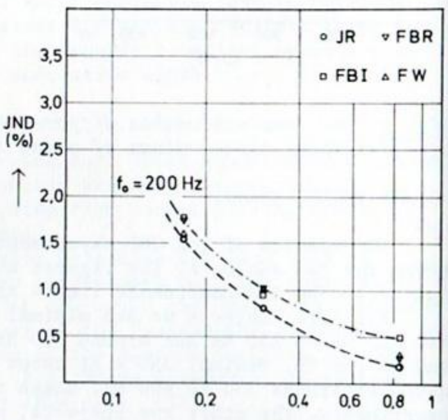

Fig. 4. The just noticeable difference in pitch (JND) of COMB-noise as a function of g for f_o=200 Hz (τ=5 ms) and four subjects

subject JR and 0.8% for subject FB. Comparing these values with the values actually measured, viz. 0.75% for subject JR and 1% for subject FB, there appears to be a good correspondence.

Further, it is interesting to compare the JND-data of COMB-noise with the JND of a periodic pulse (PP). Recent measurements in our lab. (Wittkämper, unpublished results) using exactly the same experimental procedure resulted in an average JND of about 0.3% for f_o=200 Hz. Realizing that a PP has a line spectrum, while COMB-noise has a continuous spectrum (the lines have, in fact, undergone a certain broadening), we conclude that COMB-noise with g=0.85 has "optimum sharpness of peaks", i.e. increasing the steepness of the spectral peaks by increasing g above 0.85 does not improve the JND in pitch.

7. CONCLUSIONS

- The COMB-JND is independent of f_o and equal to 0.3% on the average, if g=0.85
- The MPS-JND is dependent on f_o. Minimum values of 0.9% on the average were obtained for f_o about equal to 250 Hz and 500 Hz. This is attributed to the dominance phenomenon.
- The MPS-JND can be predicted from the COMB-JND by considering the modulation depth $M_{int.}$ of the (psychophysically measured) internal spectra of the signals.
- The internal modulation depth $M_{int.}$ of (dichotic) MPS-noise with g=0.8 is comparable to $M_{int.}$ of (monotic) COMB-noise with a g-factor of about 0.4
- For g=0.85 the COMB-JND is equal to the PP-JND.

Acknowledgements.
The MPS-JND-measurements formed part of a master-thesis by J.W. Müller.
The COMB-JND-measurements formed part of a master-thesis by F.W. Wittkämper.

REFERENCES

Bilsen, F.A., and Raatgever, J. (1973). Spectral dominance in binaural lateralization. *Acustica* 28, 131-132.

Bilsen, F.A., and Goldstein, J.L. (1974). Pitch of dichotically delayed noise and its possible spectral basis. *J.Acoust.Soc.Am.* 55, 292-296.

Bilsen, F.A. (1976). Pronounced binaural pitch phenomenon. *J.Acoust.Soc.Am.* 59, 467-468.

Bilsen, F.A. (1977). Pitch of noise signals: Evidence for a "central spectrum". *J.Acoust.Soc.Am.* 61, 150-161.

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

Houtgast, T. (1977). Auditory filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker. *J.Acoust.Soc. Am.* 62, 409-415.

Houtsma, A.J.M., and Goldstein, J.L. (1972). The central origin of the pitch of complex tones: evidence from musical interval recognition. *J.Acoust.Soc.Am.* 51, 520-529.

Raatgever, J. and Bilsen, F.A. (1977). Lateralization and dichotic pitch as a result of spectral pattern recognition. In: *Psychophysics and physiology of hearing* (E.F. Evans and J.P. Wilson Eds.), pp. 443-453, London, Academic Press.

Raatgever, J. (1980). On the binaural processing of stimuli with different interaural phase relations. *Doctoral dissertation*, Delft University of Technology.

Terhardt, E. (1972). Zur Tonhöhewahrnehmung von Klängen. *Acustica* 26, 173-199.

Wightman, F.L. (1973). The pattern-transformation model of pitch. *J.Acoust.Soc.Am.* 54, 407-416.

ADDENDUM by RAATGEVER and BILSEN:

For a better understanding of the (idealized) central-spectrum patterns of comb-noise and MPS noise, mathematical descriptions are presented here.

COMB-noise is a monotic stimulus. It is the result of white noise passing a delay-line (delay τ) with feed-back of a signal fraction g. It can therefore be characterized in the spectral domain by:

$$A_C(\omega) = \frac{\exp -j\omega\tau}{1-g \exp -j\omega\tau} . \tag{1}$$

So the corresponding power-spectrum is given by:

$$|A_C(\omega)|^2 = \frac{1}{1-2g \cos\omega\tau+g^2} . \tag{2}$$

It has maxima: $\frac{1}{(1-g)^2}$ if $\cos\omega\tau=1$ or $\omega=n\frac{2\pi}{\tau}$

and minima: $\frac{1}{(1+g)^2}$ if $\cos\omega\tau=-1$ or $\omega=(2n+1)\frac{\pi}{\tau}$.

MPS-noise is a dichotic stimulus. The spectral pattern in the very centre of the binaural activity pattern corresponds to the dichotic pitch image that is lateralized in the centre of the head. It is the result of the interaction of the (undelayed) information from both ears. The signal at one ear is the result of white noise passing a delay-line (delay τ) with feed-back of a signal fraction g. To the output is also added a negative fraction $-g/(1-g^2)$ of the input signal. The signal at the contralateral ear consists of the input noise signal multiplicated by a factor $1/(1-g^2)$. Do we suppose the interaction process to be an addition, then the central spectrum can be characterized by:

$$A_M(\omega) = \frac{1+\exp -j\omega\tau}{(1+g)(1-g \exp -j\omega\tau)} . \tag{3}$$

Waar is τ_i?

The corresponding power-spectrum is then given by:

$\tau_i = 0$

$$|A_M(\omega)|^2 = \frac{2+2 \cos\omega\tau}{(1+g)^2(1+g^2-2g \cos\omega\tau)} . \tag{4}$$

It has maxima: $\frac{4}{(1+g)^2(1-g)^2}$ if $\cos\omega\tau=1$ or $\omega=n\frac{2\pi}{t}$

and minima: 0 if $\cos\omega\tau=-1$ or $\omega=(2n+1)\frac{\pi}{\tau}$.

Combining eq. (2) and (4) we see that the power spectra of COMB-noise and MPS-noise are related as follows:

$$g|A_M(\omega)|^2 + \frac{1}{(1+g)^2} = |A_C(\omega)|^2 . \tag{5}$$

In other words: The power-spectrum of COMB-noise is a factor g less modulated and lifted over a constant factor $1/(1+g)^2$ with respect to the corresponding power-spectrum of MPS-noise.