Proceedings of the 10th International Symposium on Hearing

ADVANCES IN HEARING RESEARCH

Swabian Conference Centre, Irsee, Bavaria
26 June – 1 July, 1994

Editors

G A Manley, G M Klump, C Köppl, H Fastl & H Oeckinghaus

Technische Universität München

World Scientific
Singapore • New Jersey • London • Hong Kong

What do dichotic pitch phenomena tell us about binaural hearing ?

F. A. Bilsen

Perceptual Acoustics Lab., Applied Physics Department, Delft University of Technology P.O. Box 5046, 2600 GA Delft, The Netherlands

1 Introduction

Dichotic pitch phenomena have influenced our thinking about binaural interaction significantly. For as far as dichotic pitch values were concerned, Bilsen (1972) and Bilsen and Goldstein (1974) showed that the similarity of dichotic and monotic repetition pitch with the low-pitch of normal periodic signals (Houtsma and Goldstein, 1972) requires the existence of centrally generated spectral patterns with resolved (lower) harmonics. From such spectral patterns pitch is extracted by pattern recognition (Bilsen, 1977).

Later experiments on the lateralization of dichotic-pitch images for Huggins pitch (HP), Fourcin pitch (FP) and MPS-pitch (Raatgever and Bilsen, 1977; Raatgever, 1980) revealed that these pitch images behave like pure "time images" as postulated by Hafter and Jeffress (1968) for lateralized signals. This led us to conclude that, for wide band signals, the time image is the result of spectral-pattern recognition in which frequency information is pooled across frequency for particular interaural delays. Based on the Jeffress (1948) scheme, a Central Spectrum (CS) theory of binaural processing was developed (Bilsen, 1977; Raatgever and Bilsen, 1986) that stresses both interaural cross correlation and central spectrum recognition. Recently, the related concept of "straightness" was introduced by Stern and Trahiotis (1991).

Historically, it is remarkable that dichotic repetition pitch (DRP) triggered our thinking about central spectra (Bilsen, 1972; Bilsen and Goldstein, 1974), whereas it (still) is the dichotic pitch phenomenon that is not easily explained by the CS-theory in its present form. Recently, we discovered that DRP, contrary to the other dichotic pitch phenomena mentioned above, can be lateralized by means of an interaural level difference (ILD), though with a remarkable left-right asymmetry. This will be elaborated in the following sections. First, however, we will present new results of pitch matchings to reconfirm the older results, to determine the existence region, and to exclude the possibility of trivial cross-talk in the generation of DRP.

2 Properties of dichotic repetition pitch

2.1 Confirmation and extension of earlier results

Dichotic repetition pitch (DRP) can be heard if white noise is presented through headphones to the left and right ear of an observer with an interaural time delay (ITD) in the range of about 3 ms to 20 ms. The original results of DRP-matching experiments (Bilsen, 1972; Bilsen and Goldstein, 1974) have been confirmed in the past by Warren et al. (1981) and more recently by Hartmann (1993). All authors agree that DRP is difficult to hear. Warren et al. (1981) state: "This dichotic pitch is very difficult to hear, unlike monaural or diotic echo (repetition) pitch with the same delay periods, which is heard readily even by unpracticed listeners."

Other investigators, e.g. Fourcin (1970), have denied the existence of DRP. The present author always has had the impression that DRP is indeed a faint phenomenon, but that it is the problem of "finding" the pitch image in the head (separate from the diffuse noise image) and its low salience rather than the accuracy of pitch matchings. In order to reconcile these conflicting reports we started experiments to measure the existence region more precisely than we did before, to determine the accuracy of pitch matchings, and to look for conditions in which DRP was more difficult (or more easily?!) to perceive.

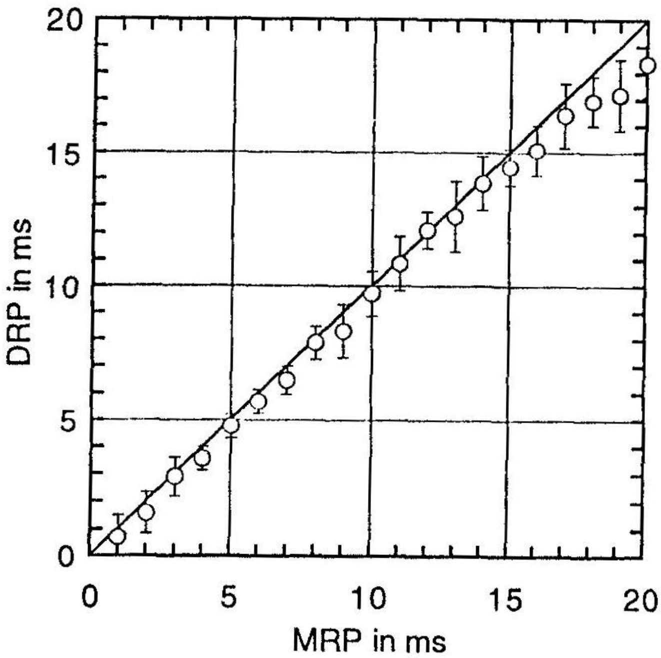

Figure 1. ITD (DRP) in ms matched against T (MRP) in ms for equal pitch.

Experiments were performed in which 3 subjects took part (For experimental details see section 4.1). They were instructed to match the pitch of a diotically presented MRP-stimulus (fixed time delay T) to a DRP by adjusting the ITD of the DRP-stimulus. They had 6 knobs to control three different in- and decrements in ITD, viz. ±1, ±0.1, and ±0.02 ms; they could switch between DRP and MRP at will. Each subject made 5 adjustments for each T (=1, 2, to 20 ms). The averaged results for all subjects are presented in Fig. 1. Similar matchings were performed with T adjustable and ITD fixed. These were judged less comfortable by the subjects, but appeared to give similar results. The solid line has been drawn through the origin under 45 degrees of arc. Though there seems to be a small deviation towards lower ITD values, the relation DRP=MRP appears to be confirmed.

The relative standard deviation, s(ΔITD)/ITD, from 15 matchings for each measured point is displayed in Fig. 2. The results clearly show that DRP can be matched with reasonable precision. The higher boundary of the existence region of DRP seems well determined by the steep decrease in matching accuracy for ITD=2 ms. The lower boundary is more diffuse. In fact DRP gradually changes into a periodicity sensation, called Infrapitch by Warren et al. (1981).

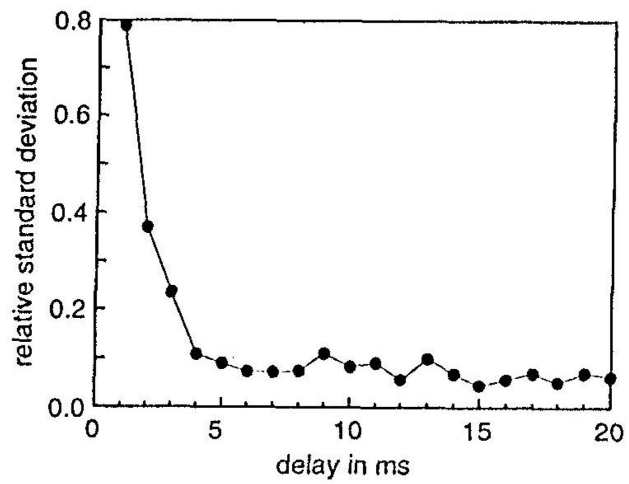

Figure 2. The relative standard deviation derived from the results of figure 1.

2.2 Arguments against cross-talk

It has been argued that DRP might be caused by a trivial form of interaural cross talk, e.g. bone conduction or air conduction, because: a) it sounds very similar to its monotic counter part (monotic repetition pitch); it has the same low pitch and the same timbre, b) it is probably the faintest dichotic pitch known, c) it behaves different from other dichotic pitch phenomena. One could imagine that the (un)delayed noise is fed to the contralateral ear by bone conduction or by cross-talk through a leaking headphone, thus producing a faint MRP in the left and/or right ear.

The following arguments and results of pilot experiments, however, exclude the possibility of an artefact by trivial cross-talk:

1) The DRP can be perceived and matched rather accurately at low over-all sensation levels. Even at a sensation level as low as 5 dB our subjects were able to match DRP to MRP. Such a result would be explainable by cross-talk only if cross-talk would be unrealistically strong. In audiology it is common practice to reckon with a rather low interaural cross-talk (40 to 50 dB),

2) If the signal to the left or right ear is attenuated with respect to the contralateral ear,

cross-talk, if effective, might result in different MRP's at the left and right ear. Attenuation of the undelayed signal might give rise to an MRP in the corresponding ear because the (delayed) signal from the contralateral ear will interfere optimally with the attenuated (undelayed) signal. In that case, however, one would expect a pitch lateralized to the undelayed ear and corresponding to ITD+ΔITD with ΔITD being the extra-interaural delay due to cross-talk. With the attenuation in the opposite channel, the pitch would correspond to ITD-ΔITD and it would be lateralized to the opposite ear. Such behaviour, however, has not been observed.

3) Possible cross-talk in the apparatus can be checked by measuring autocorrelation functions or (comb) spectra for the left and right ear signal. In our case the signals appeared to be perfectly white.

3 DRP and the central spectrum theory

The central spectrum (CS) theory was developed as a simplified analytical model for the processing of interaural delay or phase differences for low-frequency (<1500 Hz) signals. Following the Jeffress (1948) scheme of binaural interaction, filtered cochlear signals from one ear are added along internal (neural) tapped-delay lines to the undelayed signals from corresponding cochlear outputs of the contralateral ear. After squaring, a continuum arises of power (mimicking neural activity) versus frequency and internal delay, the so-called central activity pattern (CAP).

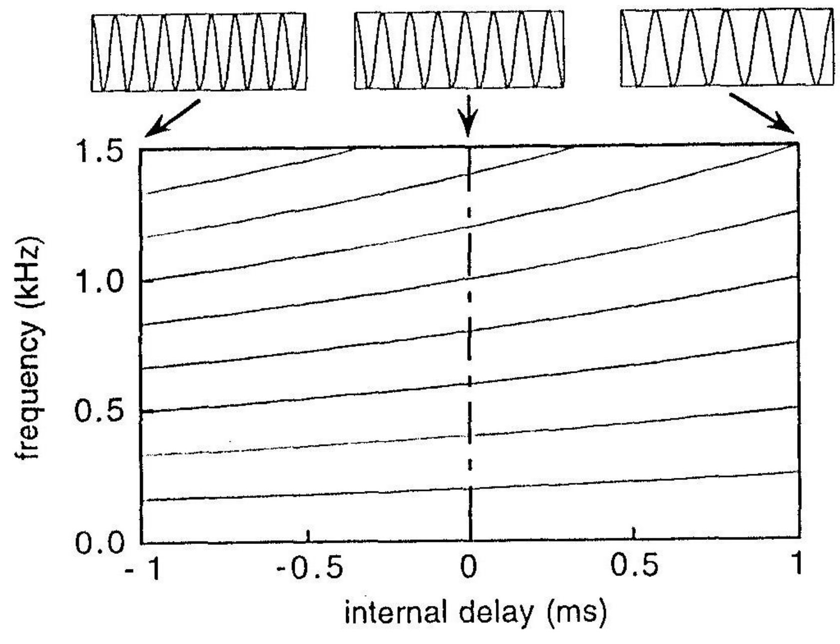

Figure 3. The CAP of DRP for ITD=T= 5ms. Only lines of maximum activity (P=2) have been plotted. At the top of the figure, central spectra are given for three different cross

For a dichotic signal with an interaural phase function $\Phi(f)$, the CAP is shown (Raatgever and Bilsen, 1986) to be equal to

$$P(f,\tau_i) = 1 + \cos\{\Phi(f) - 2\pi f\tau_i\} \qquad (1)$$

where f is frequency and τ_i is the internal delay along the (cross-correlation) time axis. For a DRP-signal with ITD = T the CAP takes the form:

$$P(f,\tau_i) = 1 + \cos\{2\pi f(T - \tau_i\} \qquad (2)$$

As an example of the CAP illustrative for DRP, curves of maximum activity (P=2) have been plotted in Fig. 3 for T=5 ms.

In the CS-theory, the central spectrum corresponding to the perceived pitch is given by the cross-section of the CAP at the internal delay τ_i =0 (dash-dotted line) which corresponds to the perceived position of the DRP-image. From Eq.(2) and Fig. 3 it might be clear that he DRP-value is correctly predicted. However, one should wonder why all the other candidate harmonic spectra at the other internal delays cannot be heard. Is the DRP the result of another mechanism, or is DRP the result of averaging central spectra across internal delay?

4 Lateralization of DRP

4.1 Experimental set-up

DRP-stimuli derived from lowpass-filtered gaussian noise with a high-cutoff frequency of 2300 Hz, were generated in a digital signal processor (Loughborough DSP 96002; sample frequency 25 kHz; (program) code generated by a Comdisco SPW system). Thus a dichotic stimulus (see simplified block diagram in Fig. 4) with an ITD and

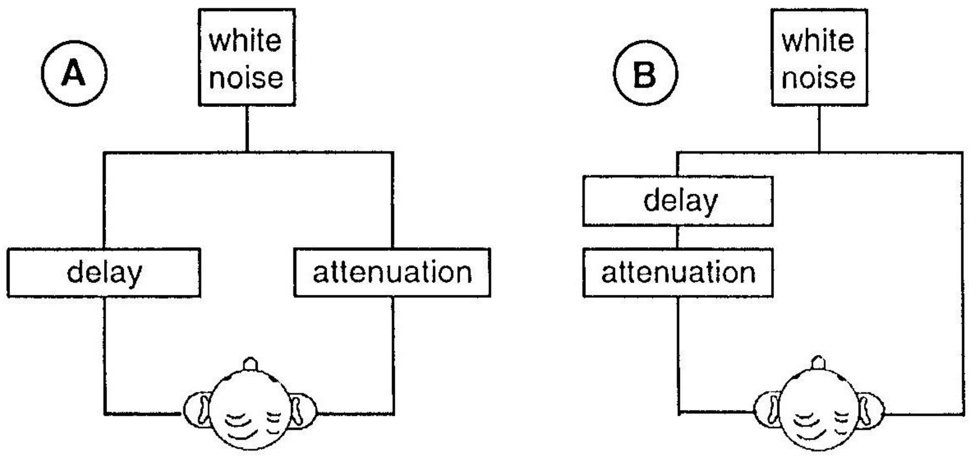

Figure 4. Block schematics of the DRP lateralization experiments; configuration A and B.

an ILD in either the same (configuration B) or the opposite-ear signal channel (configuration A) was obtained.

In order to have a comfortable perception of DRP the ITD was programmed such that an 8 note equally-tempered musical scale was played ranging from C (130.8 Hz; ITD=7.65 ms) to c (261.6 Hz; ITD=3.82 ms) with a stimulus duration of 1 second for each note; the noise itself was not switched on and off but was present continuously. The musical scale was played repeatedly, each time with a particular ILD in a range of 0 to 20 dB. For each ILD, the subject, seated in a sound proof booth and listening with headphones (Beyer DT 770) at a sensation level of about 40 dB, had to indicate the in-head position of the DRP-melody image on a linear scale ranging from -5 (left ear position) to +5 (right ear position) with 0 center position (compare Blauert, 1983, for standard lateralization procedures).

4.2 Results

Two experienced subjects participated in this experiment, in a total of four separate sessions for each subject: two sessions with configuration A of which one with left and right headphones interchanged to average out a possible bias in perceived lateralization, and the same for configuration B. Their averaged results are presented in Fig. 5; thus each measured point is the average of four observations.

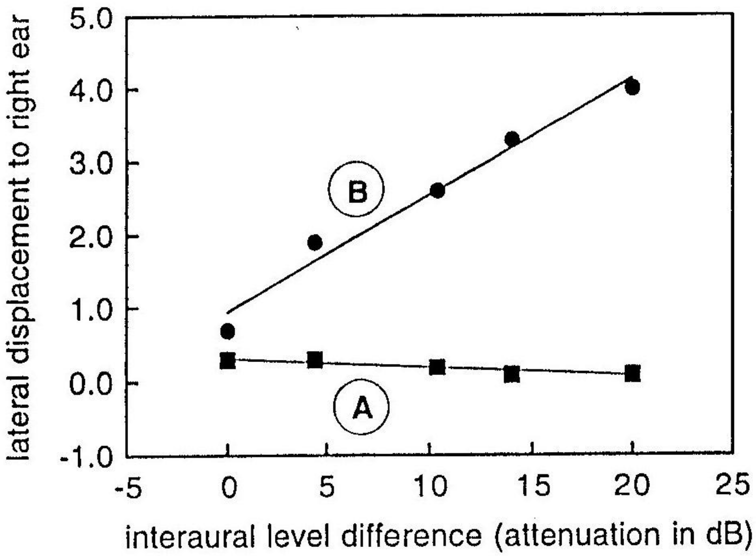

Figure 5. DRP-lateralization results for configurations A and B.

It was reported by the subjects that the DRP-image did not move within a scale, thus it was confirmed that the value of T seems to have no influence on the position of DRP. They also reported that the DRP was experienced more salient for ILD 10 dB than for 0 dB.

4.3 Discussion

The results of Fig. 5 are very remarkable because of their asymmetry. Configuration A hardly shows any shift in DRP-position with changing ILD, whereas configuration B shows a substantial shift with changing ILD. In the latter case, the relation between perceived position and ILD is about comparable to the function one observes for the lateralization of diotic noise with changing ILD (Blauert, 1983). A simple explanation for this asymmetry could not be found.

An explanation by the advocate of the devil stressing cross-talk clearly fails. If simple cross talk (by air or bone conduction) were to play a significant role, one would expect the strongest DRP in the ear provided with the attenuated signal, because the intensity of delayed and undelayed signal are more equal than in the unattenuated ear, thus providing the best-modulated comb spectrum and consequently the strongest MRP in the attenuated ear. In case B one thus would expect lateralization towards the attenuated ear However, one observes the opposite.

5 General discussion

Dichotic repetition pitch (DRP) is a faint binaural pitch phenomenon, evoked by the interaural delay of a single (white) noise signal. It appears not to be an artefact due to trivial interaural cross-talk. It has been shown in the past that its pitch behaviour provides evidence for a central-spectrum processing of pitch, rather than cross-correlation processing for which (according to Fourcin) two uncorrelated noise sources would be required. Nevertheless, as discussed in section 3, DRP-extraction is still not fully understood. The present experiments were intended to get further understanding of the binaural processes involved.

The present results on the accuracy of DRP matchings show that its existence region is sharply limited to ITD's larger than 2 ms, whereas the lower border is more diffuse and extends into the region of rattle- or infra-pitch. This is not unexpected since the lateralization of binaural signals in general is limited to ITD's smaller than 1 ms. Note that the CS-concept correspondingly has no need for (neural) internal delays longer than say 1 ms.

In former papers, it has been shown that the dichotic-pitch images of HP, FP, and MPS can be lateralized by an (extra) ITD. The lateralization shows great similarity with the behaviour of the time image. ILD's, on the contrary, are ineffective in lateralizing these dichotic pitches; they only affect their salience. In the present paper, it is shown that, of course, an extra ITD does not influence the lateralization of DRP (How would the binaural system, for a change in ITD, be able to discriminate between a change in pitch and a change in lateralization?). However, for DRP, ILD's are effective in lateralizing the pitch image, be it in an asymmetric way and with an unexpected increase in salience for ILD's in the order of 10 dB.

The final tempting conclusion, thus, should be that there are two separate mechanisms: a mechanism to evaluate ITD's and a mechanism to evaluate ILD's. DRP could very well be the product of the ILD-mechanism exclusively, contrary to HP, FP and

MPS that are clear exponents of the ITD-mechanism.

6 Acknowledgements

The experiments described in section 2 were performed by Han Dols in fulfilment of his master's thesis. Thanks go to the subjects JR, HD and FB for their persistence.

7 References

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. (1972) Pitch of dichotically delayed noise. In: E. de Boer, B.L. Cardozo and R. Plomp (Eds.), Hearing Theory 1972, I.P.O., Eindhoven, the Netherlands, pp. 5-8.

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.

Blauert, J. (1983) Spatial Hearing. MIT Press, Cambridge, U.S.A..

Cramer, E.M. and Huggins, W.H. (1958) Creation of pitch through binaural interaction. J. Acoust. Soc. Am. 30, 413-417.

Fourcin, A.J. (1970) Central pitch and auditory lateralization. In: R. Plomp and G.F. Smoorenburg (Eds.), Frequency Analysis and Periodicity Detection in Hearing, Sijthoff, the Netherlands, pp. 319-328.

Hafter, E.R. and Jeffress, L.A. (1968) Two image lateralization of tones and clicks. J. Acoust. Soc. Am. 44, 563-569.

Hartmann, W.M. (1993) Personal communication.

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.

Jeffress, L.A. (1948) A place theory of sound localization. J. Comp. Physiol. 41, 35-39.

Klein, M.A. and Hartmann, W.M. (1981) Binaural edge pitch. J. Acoust. Soc. Am. 70, 51-61.

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

Raatgever, J. (1980) Binaural time processing and time-intensity trading. In: G. van den Brink and F.A. Bilsen (Eds.), Psychophysics, Physiology and Behavioural Studies in Hearing, D.U.Press, Delft, the Netherlands, pp. 425-428.

Raatgever, J. and Bilsen, F.A. (1986) A central spectrum theory of binaural processing. Evidence from dichotic pitch. J. Acoust. Soc. Am. 80, 429-441.

Stern, R.M. and Trahiotis, C. (1991) The role of consistency of interaural timing over frequency in binaural lateralization. In: Y. Cazals, K. Horner and L. Demany (Eds.), Auditory Physiology and Perception, Pergamon, Oxford, pp.547-554.

Warren, R.M., Bashford, J.A. and Wrightson, J.M. (1981) Detection of long interaural delays for broadband noise. J. Acoust. Soc. Am. 69, 1510-1514.

Comment by J. Culling

A distinction must be made between across-frequency grouping and across-frequency integration. The former groups frequency channels sharing common features which are likely to be associated with a common sound-source. The latter gathers channels together indiscriminately. Stern & Trahiotis (1991) demonstrated across-frequency integration of interaural timing information in lateralization, but Culling and Summerfield (1994) found no evidence of across-frequency *grouping* by common interaural delay in a speech identification task. Conventional models of binarual masking release simply detect frequency channels which display interaural decorrelation and can explain Huggins-pitch, MPS-pitch and binaural edge pitch without using across-frequency grouping. Why, therefore, is it necessary to invoke across-frequency grouping (as described in the spoken presentation) in order to explain these phenomena?

Culling, J.F. and Summerfield, Q. (1994) Proceedings of 127th Meeting of the Acoustical Society of America.

Reply

I agree that conventional models like the EC-model can explain Huggins pitch, or generally: the generation of "central spectrum peaks". But, in order to correctly predict both the low-pitch and its (lateralized) image of e.g. MPS, we postulated a mechanism that "takes together" (groups, not integrates) the CS-peaks at the proper common internal delay. We did not study or encounter dichotic low-pitch phenomena inconsistent with that view so far.