Proceedings of the 12th International Symposium on Hearing

Physiological and Psychophysical Bases of Auditory Function

Carlton de Brug Resort, Mierlo, The Netherlands
4-9th August 2000

Editors

D. J. Breebaart
A. J. M. Houtsma
A. Kohlrausch
V. F. Prijs
R. Schoonhoven

The case of the missing central spectra

F. A. Bilsen

Applied Physics Department, Delft University of Technology
P.O. Box 5046, 2600 GA Delft, The Netherlands
bilsen@akphp.tn.tudelft.nl

1. Introduction

This paper deals with dichotic pitch phenomena, the Dichotic Repetition Pitch (DRP) in particular. Dichotic pitch phenomena, in general, are considered natural epiphenomena of the process of binaural hearing, and it is expected that the study of dichotic pitches can be successfully applied to increase our knowledge about binaural hearing and to model ecologically-important functioning like sound localization and signal detection in noise (Stern and Trahiotis, 1995). Several dichotic pitch phenomena are known: the Huggins Pitch (Cramer and Huggins, 1958), the Fourcin Pitch (Fourcin, 1962), the Dichotic Repetition Pitch (Bilsen, 1972), the Multiple Phase Shift Pitch (Bilsen, 1976), and the Binaural Edge Pitch (Klein and Hartmann, 1981).

DRP can be perceived with stimuli consisting of a single interaurally-delayed white noise, with delays typically in the range of 4 to 20 ms. Reports on the existence of DRP came from Bilsen (1972, 1995), Bilsen and Goldstein (1974), Warren et al. (1981), and Culling (1996). It was not reported by Blodgett et al. (1956) though they studied the sidedness of interaurally-delayed noise with relatively large interaural delays. Further, its existence was denied by Fourcin (1962, 1970) and disputed by Hartmann (1996). In contrast, a rather strong sensation of dichotic pitch is observed if two or three independent interaurally-delayed noises are presented simultaneously: the Fourcin Pitch (Fourcin, 1962, 1970; Bilsen and Goldstein, 1974; Bilsen and Wesdorp, 1974; Bilsen, 1977; Culling et al. (1998); Bilsen and Raatgever, 2000).

Based on Jeffress' physiological binaural network (Jeffress, 1948), the Central Spectrum theory (Bilsen, 1977; Raatgever and Bilsen, 1986; Bilsen and Raatgever, 2000) proved to provide a suitable framework for explaining both the perceived pitch value and the perceived lateralized position of the different dichotic pitch phenomena. However, the DRP though in fact having been the trigger for the theory anno 1977, formed a problem for the theory anno 1986, because a continuum of competing central spectra (and pitches) is predicted while only a single pitch with a single centralized image had been reported.

Recently, we (the present author, during a series of experiments together with W.M. Hartmann and J. Raatgever (Hartmann et al., 1998)) experienced that with proper dynamic stimulus presentation, DRP seemed to give up its unique value and its centralized image position. A deviating pitch at a different position appeared to be perceivable. In the present contribution, we show that it is possible to measure systematically, for a fixed value of the interaural delay, different pitches with each pitch percept located at a particular lateralized position.

In the following sections, the CS model is summarised and the predictions for DRP presented, new data on DRP are reported, and existing data on DRP are reconsidered. The contribution ends with general implications for binaural hearing theory.

2. DRP and its prediction by the Central Spectrum theory

Most models dealing with binaural hearing assume a scheme of interaural signal processing based on principles first described by Jeffress (1948). In this scheme, the left and right ear signals undergo a peripheral frequency analysis followed by a process of cross correlation in a network consisting of tapped parallel (neural) delay lines. In such a network, a two-dimensional activity pattern is generated that shows neural activity as a function of internal delay and frequency. Lateralization on the basis of interaural time differences can be explained by a cross-correlation process.

The Central Spectrum model (Raatgever and Bilsen, 1986), in essence, based on Jeffress' network, predicts dichotic pitch values as well as their intracranial positions. It calculates a Central Activity Pattern (CAP) as a function of frequency and internal delay. A central pitch processor scans this CAP for a sharp isolated peak giving rise to a pure-tone-like pitch, or a well modulated periodic spectral pattern at a particular internal delay giving rise to a "low" pitch (synonymously: residue pitch, virtual pitch). This pattern selected, the "Central Spectrum" (CS) (compare Bilsen, 1977), is claimed to predict the value of the pitch, while the corresponding internal delay determines the perceived lateral position of the pitch image (Stern and Trahiotis (1995) use the term 'straightness' for the consistency of the internal delay over frequency for lateralized stimuli, in general).

The selection mechanism for a Central Spectrum to be a serious candidate as predictor of low pitch, was not mathematically specified in the original formulation of the model (Raatgever and Bilsen, 1986). The only essential criteria for the result of the scanning process as described above, to be a CS candidate are assumed to be: 1) periodicity of the peak distance measured along the frequency axis, for one and the same internal delay, and 2) a large modulation depth in the spectrum selected.

Using Eq. (6) from Raatgever and Bilsen (1986), the Central Activity Pattern (CAP) of a DRP stimulus follows as

$$CAP(f, \tau_i) = 1 + \cos(2\pi fT + 2\pi f\tau_i). \tag{1}$$

Note that for each value of the internal delay τ_i a well-modulated cosinusoidal function of frequency is found waxing and waning between 0 and 2, thus with a peak-to-valley ratio equal to infinity. A pitch equal to $\{1/T\}$ is predicted for $\tau_i=0$, but other pitches equal to $\{1/(T+\tau_i)\}$ are expected at all possible values of τ_i in $CAP(f, \tau_i)$ (see Fig. 1).

This implies that no pitch at all would be predicted to be perceivable due to mutual competition of an infinite number of candidate spectra The historical reports of only a single faint pitch corresponding to $\{1/T\}$ positioned in the center of the head ($\tau_i = 0$) are explained by the CS model only if strong prevalence for $\tau_i = 0$ would be assumed. Other data on dichotic pitch, but also on lateralization with conventional stimuli plead against such an assumption.
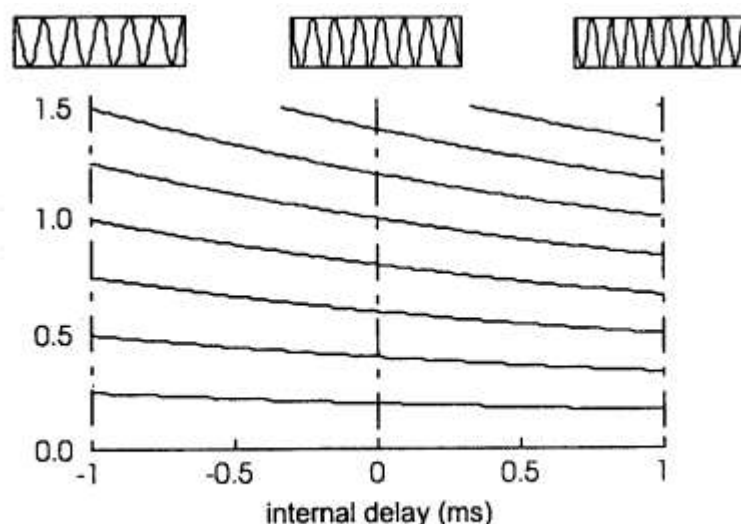
Figure 1. Schematics of the Central Activity Pattern (CAP) for a DRP stimulus with T=5 ms. The internal delay (in ms) is plotted along the horizontal axis. The frequency (in kHz) is plotted along the vertical axis. Only lines of maximum activity (CAP=2) have been plotted. At the top of the figure, cosinusoidal Central Spectra resulting from cross sections are given for internal delays -1, 0, and 1 ms.

3. Pitch values and pitch image positions of DRP for a fixed T

A continuous DRP signal with two continuously alternating values of the interaural delay, {T} and {$T+\Delta T$}, was processed in a digital signal processor (Loughborough DSP 96002) implemented in a host-PC configuration. A sampling frequency of 20 kHz was used and the (program) code was generated by SPW software (ALTA Comdisco). The original noise was derived from an analog gaussian noise source, lowpass filtered with a cutoff frequency of 5 kHz. The alternation of the two interaural delay values was chosen at a rate of 0.33 Hz, thus each note of the DRP interval had a duration of 300 ms. The stimuli were presented by headphones (Beyer DT 770) at a sensation level of 40 dB. The subjects were seated in a double-wall sound proof booth.

A match stimulus was used consisting of a periodic impulse of 100-μs duration and repetition frequency F (in Hz), that served as a lateralization pointer and as a pitch comparator simultaneously. Its lateralized position (compare internal delay) indicated by the interaural delay τ (in ms) could be manipulated by the subject directly by turning the 0.1-ms step knob of a delay line (Allen Avionics) either in the left or right ear signal path. The pitch value (compare F) was controlled by the subject turning the knob (scale invisible) of an RC generator in conjunction with an impulse generator. The match stimulus could be switched on and off by the subject at will.

Subjects were instructed to match both the pitch value and the pitch image position of the highest note of the DRP interval by adjusting F and τ for each DRP percept they could "discover" in the perceptual space filled with noise. They were encouraged to look for percepts lateralized to either the left or right ear. After a match of both pitch and lateralized position was made to the subject's satisfaction, F was

read from a frequency counter and τ was read from the 0.1-ms (or 1.0 ms) knob position on the delay-line front panel.

Three normal-hearing subjects took part in the pilot experiments. All three had no difficulty finding and matching a centralized ($\tau = 0$) percept with a pitch $F = 1/T$, i.e. the classical DRP percept. Only one subject (FB) was able to direct his attention to other percepts at different places.
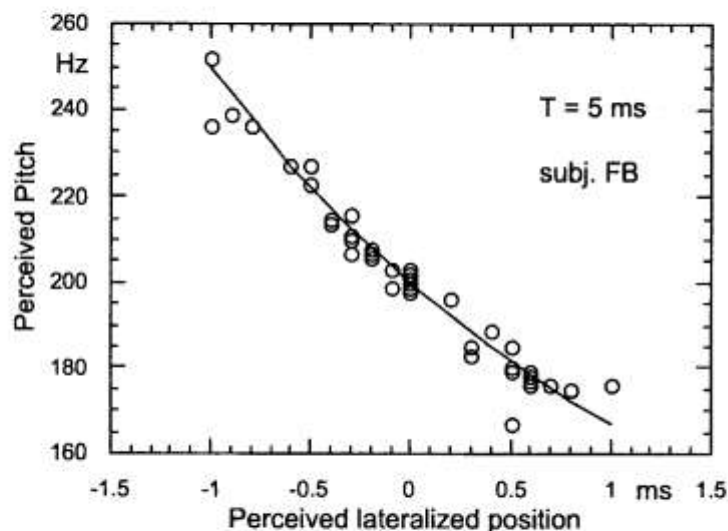

Figure 2. Individual matches (open circles) of pitch and pitch-image position by subject FB for a DRP stimulus with $T = 5$ ms. The value of the pitch (expressed as the fundamental frequency F of a periodic-pulse matching signal in Hz) is plotted along the vertical axis. The value of the perceived position of the pitch-image (expressed as the interaural delay τ of the periodic pulse signal in ms) is plotted along the horizontal axis. The solid line represents the prediction by the Central Spectrum model.

During 5 sessions of about one hour in length, subject FB produced matches for T equal to 5 ms (with $\Delta T = 0.5$ ms). Individual pitch with pitch position matches are presented in Fig. 1 by open circles. The solid line represents the prediction by the CS model according to Eq. (1). Except a few, the measured points closely fit the prediction. There appear to be three values of the perceived position (compare τ) where a relatively large number of matches was made, i.e. -0.2, and 0.6 ms, and especially 0 ms (the classical DRP). It should be remarked that subject FB tried very hard to find percepts for the whole range of possible lateralized positions, but the values mentioned appeared to stand out as positions with a certain preference.

To investigate the possible influence of an interaural intensity difference (IID) on the value of the pitch and/or its lateralized position, matchings were performed for the above stimulus configuration with an extra IID. The results are given in Fig. 2.

Open circles represent matches for an attenuation equal to 10 dB applied to the left ear channel; open triangles similarly matches for the right ear channel. The solid line again is the CS-model prediction following Eq. (1). Although the fit seems less perfect compared to the case without IID, also these matches are considered to support the theory. Further, no IID effect is observed because, as the circles and triangles seem not to behave differently.
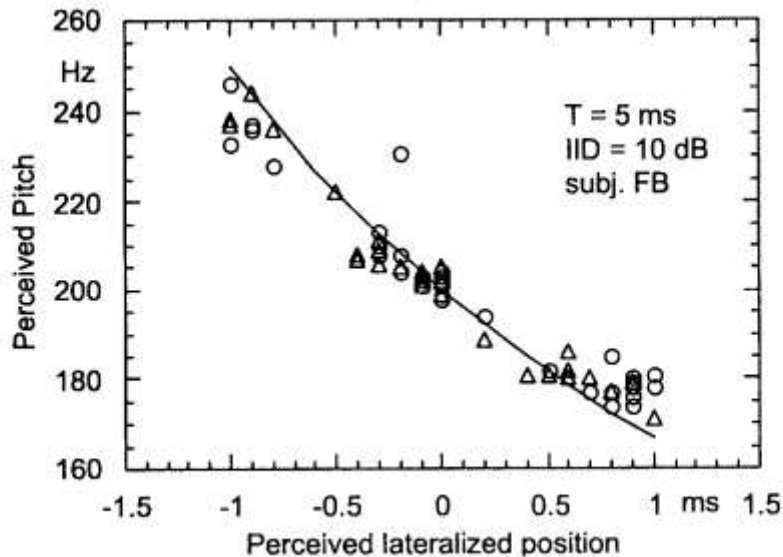
The case of the missing central spectra

Figure 3. Perceived pitch and pitch-image position by subject FB for a DRP stimulus with T=5 ms. Open circles represent individual matches for an attenuation of 10 dB applied to the left ear. Open triangles idem for an attenuation of 10 dB applied to the right ear. Further, see legend of figure 1.

4. Discussion

The present measurements by subject FB may be considered as evidence for the existence of a continuum of possible pitches as predicted by the CS model following Eq. (1). The process of "finding" different DRP percepts seems best described as trying hard to concentrate on a particular point of the lateralization axis. Usually, after some time (order of magnitude 10 sec.), such a percept "pops out" of the noise filling the head. Giving the match stimulus a pre-set lateralization was expected, but did not turn out to be helpful in catching a DRP percept at a particular place.

The reason that the present result was obtained only recently, obviously was the present way of presentation of the DRP stimulus, viz. a continuous alternation of two interaural delays $(T+\Delta T)$ instead of only one fixed T, as done mostly during our experiments in the past. One might speculate about the nature of the underlying processes. Does dynamic presentation cause a facilitation or rather a released inhibition of the central spectrum patterns. Maybe, it is only a matter of increased attention directed to these alternative patterns. Also it is not clear why only one subject out of three could perform the present task.

It can be concluded that, at least for one subject, the "missing central spectra" were found. It might be interesting to realise that the present result cannot be predicted from the cross correlation function alone. Of course not, as the perceived pitches are generated by the specific structure of the binaural network itself. The fact that these pitches obey the rules of the Central Spectrum theory, i.e. that their central spectra follow from cross sections with a common internal delay and not from another type of cross section most probably is indicative for the structure of the neural network after the "Jeffress' stage". On the other hand, for the DRP signal at least, other type cross sections through the CAP (e.g. not perpendicular to the internal delay axis; compare Fig. 1) never will display a strictly harmonic spectrum.

Other dichotic pitch phenomena like the MPS pitch or the Fourcin Pitch, obey the common-internal-delay property of the CS model also (if, for some configurations, small deviations in internal delay at the dominant frequency region are allowed) (Bilsen and Raatgever, 2000). They evoke rather clear percepts at (un)ambiguous lateralized positions, but there the stimulus itself basically conveys the information needed for the specific internal delay found. Therefore, in general, these stimuli do not exclude the possibility that information from the CAP for different interaural delays be combined (Culling et al., 1995; Bilsen et al., 1998).

Culling et al. (1998) proposed a modified Equalization-Cancellation (mE-C) model to account for dichotic pitches as illusions of binaural unmasking. Essentially, their model performs an equalization by adjustment of internal delay (and/or level) in each frequency channel (auditory filter) independently. Apparently, many different delays may be involved depending on the dichotic stimulus, instead of the same for all frequency channels as in Durlach's original concept. For the DRP stimulus, the mE-C model does not produce any "recovered spectrum" at all. This non-prediction might be considered at odds with the possible existence of a continuum of pitches as predicted by Eq. (1). On the other hand, it might as well be considered indicative for the extremely low salience of DRP, or the difficulty that some listeners have to perceive or match DRP.

An IID of 10 dB does not seem to affect the data at all (compare Fig. 1 and Fig. 2). Bilsen (1995) measured the lateralized position of DRP (played as an 8-note equally-tempered musical scale) for two configurations, A) with the attenuation and the delay in the signal paths to opposite ears, and B) with the attenuation in the same signal path as the delay. In agreement with the present results, case A did not show substantial IID sensitivity, however case B did. At present, no explanation can be given for this asymmetry. Subject FB remembers that the 8-note scale became mistuned with increasing IID, but this observation was not seriously considered. Reconsidering it now, it might be explained by an IID-induced shift of attention in the CAP to different internal delays. The present IID insensitivity is in accordance with older and new results on the Fourcin Pitch, MPS and Huggins Pitch (Bilsen and Raaatgever, 2000), but partly in contradiction with recent experiments by Grange and Trahiotis (1996) on Huggins Pitch.

It is tempting to hypothesize that the processing of ITDs and IIDs in general (thus also for ecological stimuli) takes place beyond the level at which dichotic pitches are generated. Thereby, it seems essential that IIDs only affect the intracranial position of the percept of stimuli that have "bi-monaural" or "bi-monotic" existence, i.e. of which two near-copies exist at the level of the left and right cochlea. Such stimuli essentially can give rise to two near-identical monotic percepts also, if presented either at the left or right ear alone. A dichotic pitch is unique insofar that it has no near-copy, neither at the cochlear level nor somewhere else at the same level of processing. As a consequence its intracranial position is not essentially affected by IIDs (see also Bilsen and Raatgever, 2000).

5. References

Bilsen, F.A. (1972) Pitch of dichotically delayed noise. In: B. L. Cardozo (Ed.), *Hearing Theory 1972.* I. P. O., Eindhoven), 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.

Bilsen, F.A. (1995) What do dichotic pitch phenomena tell us about binaural hearing. In: G. A. Manley, G. M. Klump, C. Köppl, and H. Fastl (Eds.), *Advances in Hearing Research*. World Scientific, London, pp. 334-341.

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., and Raatgever, J. (2000) On the dichotic pitch of simultaneously-presented interaurally-delayed white noises. Implications for binaural theory. J.Acoust. Soc. Am. **108**, 272-284.

Bilsen, F.A., Meulen, A.P. van der, and Raatgever, J. (1998). Salience and JND of pitch for dichotic noise stimuli with scattered harmonics: grouping and the central spectrum theory. In: A. R. Palmer, A. Rees, A. Q. Summerfield, and R. Meddis (Eds*), Psychophysical and Physiological Advances in Hearing*. Whurr, London, pp 403-411.

Blodgett, H.C., Wilbanks, W.A., and Jeffress, L.A. (1956) Effect of large interaural time differences upon the judgment of sidedness. J. Acoust. Soc. Am. 28, 639-643.

Culling, J.F. (1996) Dichotic pitches as illusions of binaural unmasking. J. Acoust. Soc. Am. 99, 2515.

Culling, J.F., and Summerfield, A.Q. (1995) Perceptual separation of concurrent speech sounds: absence of across-frequency grouping by common interaural delay. J. Acoust. Soc. Am. 98, 785-797.

Culling, J.F., Marshall, D.H., and Summerfield, A.Q. (1998). Dichotic pitches as illusions of binaural unmasking. II. The Fourcin pitch and the dichotic repetition pitch. J. Acoust. Soc. Am. 103, 3527-3539.

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, Leiden, pp. 319-328.

Grange, A.N., and Trahiotis, C. (1996) Lateral position of dichotic pitches can be substantially affected by interaural intensitive differences. J. Acoust. Soc. Am. 100, 1901-1904.

Hartmann, W.M. (1996) Almost dichotic repetition pitch. Michigan State Psychoacoustics Report 93, pp. 1-6.

Hartmann, W.M., Bilsen, F.A., and Raatgever, J. (1998) Dichotic pitch. Minutes of the 1998 experiment session. Michigan State Psychoacoustics Report 104, pp. 1-13.

Jeffress, L.A. (1948) A place theory of sound localization. J. Comp. Physiol. Psychol. 61, 468-486.

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

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. (1995) Models of binaural interaction. In: B. C. J. Moore (Ed.), *Hearing*. Academic, London, pp. 347-386.

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 from Akeroyd

How does the salience --- or pitch strength --- of the dichotic repetition pitch depend upon the lateralization upon which the listener is attending? Does the salience follow the form of the "central-weighting" function (e.g., Colburn, 1977): i.e, is the salience equally strong for those pitches heard at or near the center of the head and then gradually weaken as the lateralization of the pitch is increased?

Reference: Colburn, H.S. (1977). "Theory of binaural interaction based on auditory-nerve data. II. Detection of tones in noise". JASA 61 525-533.

Reply

I did not try to measure the salience of the different pitches. In the range of internal delays found, the salience did not seem to vary as a function of internal delay in a systematic way. On the other hand, for some values and especially for 0 ms internal delay, a relatively large number of matches was made, indicating a relatively

high chance for these pitch percepts to be detected. One might perhaps call these pitches more salient.

Comment by Buus

What is the role of the alternating pitches (delays) in your test stimulus? Is this pitch alternation necessary for hearing out the different pitches in the different lateral positions?

Reply

In the past, we often used a stimulus presentation with a constant external delay (T). Then only the pitch (1/T) at an internal delay of 0 ms was perceived. In the present paper, using a continuous alternation of two external delay values (T and 1.1T), also other possible pitches (1/(T+t)) at other internal delays (t) were found. Apparently, our central pitch processor (including attention) is "helped" by this way of presentation. I consider this an example of the general fact that the sense organs are more sensitive to changing stimuli than to constant stimuli.

The case of the missing central spectra