inter•noise 94
YOKOHAMA-JAPAN, AUGUST 29-31

LOCALIZATION AND DETECTION OF SIGNALS IN NOISE

F. A. Bilsen

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

subject classification 61.2, 63.3, 79.2

INTRODUCTION

A review is given on the abilities of the auditory system to localize and detect acoustic signals in silence and noisy environments. In recovering "wanted" signals from unwanted signals binaural interaction is essential. It has been well established today that three physical parameters play a major role: the head-related-transfer functions (HRTF) of the pinnae for sound incidence in the vertical and horizontal plane, interaural intensity differences (IID) and interaural time differences (ITD) for sound incidence in the horizontal plane (see Blauert, 1983, Colburn & Durlach, 1978, for extensive reviews).

Modern models of binaural interaction, dealing with these parameters, have in common a process of cross-correlation following cochlear frequency analysis (critical bands) to evaluate ITDs. They more or less deviate in their dealing with the combined presentation of ITDs and IIDs (e.g. Lindemann, 1986) and in the final process of detection of "wanted signals".

A model tat specifically deals with ITD-processing and the extraction of a wanted signal is the central-spectrum model (Bilsen, 1977; Raatgever and Bilsen, 1986). It postulates the extraction of a "central spectrum" from a central activity pattern (CAP) resulting from the cross-correlation process by pooling information across frequency, for a particular internal delay. Known spectral features such as spectral periodicity are used to select the wanted signal from the remaining (noisy) activity.

Present and future applications of head-related sound registration and binaural modeling in the field of noise control, signal detection and quality measurement, are shortly indicated in the closing section

LOCALIZATION VERSUS LATERALIZATION

Normally, we are able to localize sound sources in our environment with reasonable precision: the "localization blur" is in the order of 1 degree of arc for the horizontal plane, and 10 degrees for the vertical plane (see Blauert, 1983, for a detailed review). Physical parameters involved are the interaural intensity difference (IID), the interaural time difference (ITD) and head-related transfer functions (HRTF) (see Fig. 1a and g). Although errors are sometimes made, the perceived sound source (the "auditory event") is located in space at about the same position as the sound source (the "physical event"). Under daily-life conditions, IIDs, ITDs and HRTFs are coupled parameters.

In laboratory conditions, though, we are able to uncouple these parameters. If we present signals by headphones, the listener will generally locate the auditory event inside-the-head (IHL) instead of outside-the-head. Fig.1b sketches a traditional experimental set-up to study the effect of IID and ITD separately. Generally, the listener reports that, with a change in ITD (range: -1 to +1 ms) or a change in IID (range: -10 dB to +10 dB), the auditory event moves in the head along a line that connects the left and right ear; this is called *lateralization*. Just noticeable differences (JND) or lateralization blur are in the order of 0.01 ms and 1 dB respectively.

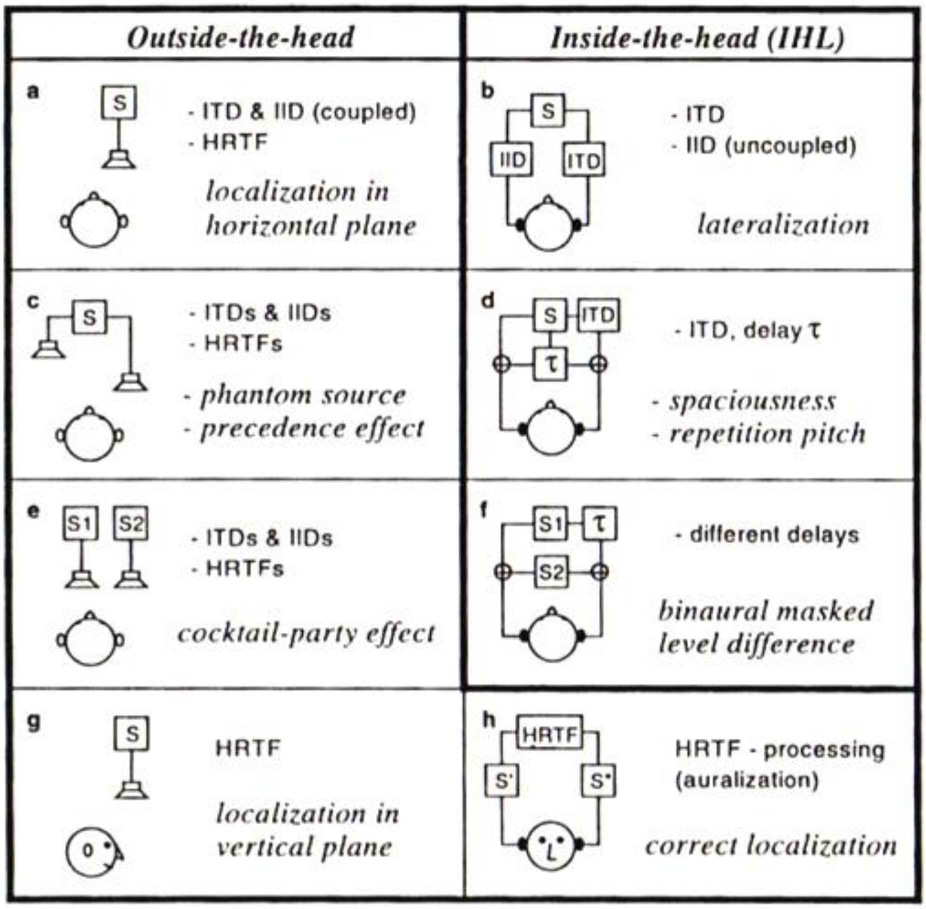

Figure 1. Perception of signals from external sources (left column), and signals reproduced by headphones (right column); S = signal, ITD = interaural time difference, IID = interaural intensity difference, HRTF = head-related transfer function. For further explanation see text.

Generally, the time-fine-structure of a signal is effective in conveying ITD-information for frequencies below 1500 Hz, with a dominant region around 500 to 600 Hz (Bilsen & Raatgever, 1973). For higher frequencies, the signal envelope is important, if the modulation frequency is lower than about 300 Hz. All (critical band filtered) signal parts play a role in IID-processing.

If a signal is presented by more than one loudspeaker (e.g. conventional stereo-set-up as given in Fig. 1c) one normally perceives the auditory event between the loudspeakers, as if there were a *phantom source* (This is called summing localization by Blauert, 1983). For time delays larger than about 1 ms between the signals, the *precedence effect* comes into play, i.e. the sound S is heard as coming from the direction of the loudspeaker of which the sound arrives first at the listener' ears. The delays involved can be as large as 50 ms for speech, but only 5 ms for sound clicks. In the laboratory situation (Fig.1d) the perception is different. Image broadening (*spaciousness*) coupled with decreased accuracy in lateralization, and coloration (*repetition pitch*), are perceived (Bilsen, 1977; Salomons et al., 1991; Potter, 1993; see also Houtgast & Aoki, 1994).

If direction-dependent HRTFs are taken into account, it is possible to obtain correct outside-the-head localization with headphone presentation (see Fig. 1h). The results are optimal if the subjects own HRTFs are used in the calculation (simulation) of the sound field (e.g. Wightman & Kistler, 1992) and if head movements are appropriately incorporated in the simulation (Bronkhorst, 1993).

DETECTION

If more than one sound signal arrive at the ears of a listener at the same time, it will be more difficult for him to perceive the wanted sound. In Fig. 1e, for example, S1 may be a wanted source, e.g. a person talking to the listener, whereas S2 may be noise from another direction in the same room. Our ability to more easily detect the wanted sound using both ears instead of one ear, is called the *cocktail-party effect*. Apparently, the binaural system uses both ITDs and IIDs as well as spectral information (HRTFs) to separate the two signals and facilitate detection.

In the laboratory situation (Fig. 1f) the corresponding release of masking has been studied extensively. If the interaural time (or phase) relations are different for signal (S1) and masker (S2) *a binaural masked level difference* (BMLD) is measured with respect to the situation where both signal and masker have the same interaural phase. The BMLD can amount to 15 dB for sine tones in noise. The effect is maximum for frequencies around 500 Hz, the dominant frequency region. Durlach developed the Equalization and Cancellation theory to account for BMLDs in a quantitative way (see Colburn & Durlach, 1978, for a review, also of other detection theories). A theory that specifically deals with the detection of wide-band (sub) signals in noise is the central spectrum theory.

CENTRAL SPECTRUM THEORY

Dichotic pitch phenomena have influenced our thinking about binaural interaction significantly. Bilsen and Goldstein (1974) have shown that the similarity of dichotic and monotic repetition pitch with the low-pitch of normal periodic signals 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 the dichotic-pitch images of 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 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 that incorporates both interaural cross-correlation and central spectrum processing (Bilsen, 1977; Raatgever and Bilsen, 1986). Recently, the related concept of "straightness" was introduced by Stern and Trahiotis (1991).

The CS-theory is 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).

In the extended CS-model, envelope information (for frequencies above 1500 Hz) undergoes a similar cross-correlation process as postulated for the low-frequency time-fine structure (Raatgever & van Keulen, 1992).

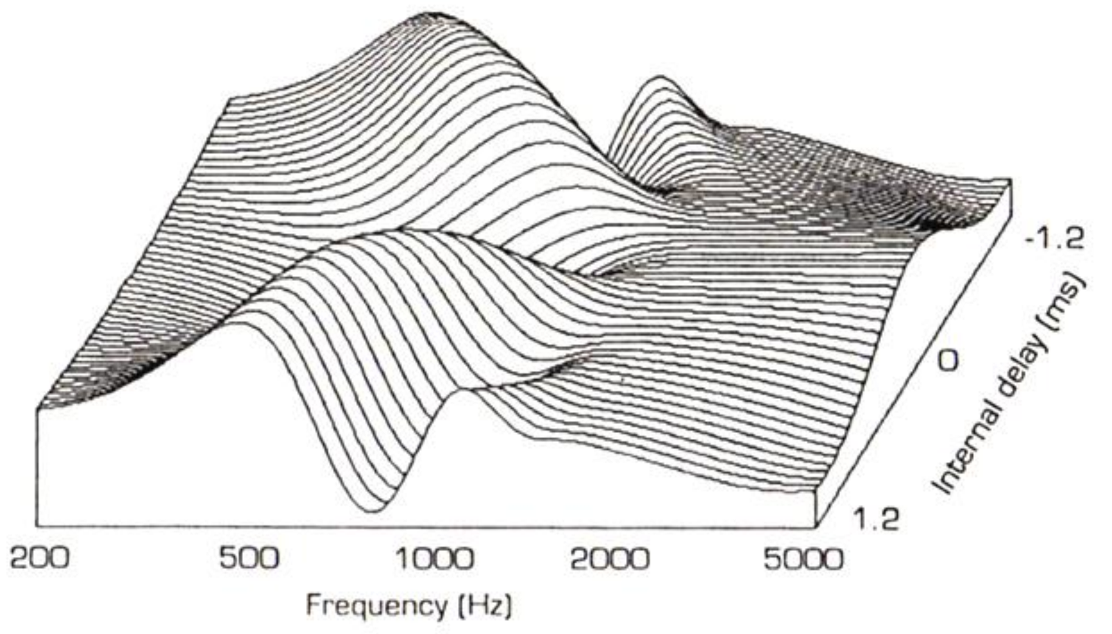

Figure 2. Central activity pattern (CAP) for antiphasic white noise

To illustrate the extended CS-model, in Fig.2 a CAP is given for antiphasic white noise. For low frequencies it shows the (diffuse) ambiguous lateralization due to time-fine structure, while at high frequencies a central ridge can be seen due to the natural signal envelopes in the (critical band filtered) noise that remain unaffected by the interaural antiphasic condition. This ridge leads to a centrally located image of the high frequencies.

During the oral presentation, examples of CAPs of dichotic-pitch stimuli will be presented that show the ability of the binaural system to select particular ("wanted") signal parts from the remaining (diffuse) noise. As such, the binaural strategy of the central spectrum theory might well be an appropriate algorithm for measurements and future applications in binaural technology.

BINAURAL TECHNOLOGY

It is well-known today that the (A-weighted) sound pressure level alone is not an adequate predictor of loudness, noise load, or annoyance. Other parameters like spectral composition, temporal structure, spatial distribution of sound sources play a significant role also.

In a study by Notbohm et al. (1992) the responses of subjects to noises recorded and presented conventionally were compared with the responses to sounds recorded by means of an artificial head (see also Genuit, 1991; Hellbrück & Kilcher, 1993). Clear evidence was found that, as far as industrial noise is concerned, spatial distribution of noise sources has a significant influence on physiological responses. Contrary to a single omni-directional microphone, the artificial head appears able to convey important perceptual aspects like direction of sound incidence, masking, movement of sound sources, etc.

Also for the study and evaluation of the sound quality of technical products in real and virtual acoustic situations, the head-related recording or calculation (simulation) of sound fields might be very useful.

A different application of binaural modeling might be the reduction of disturbing noises for hard-of-hearing peoples in a cocktail-party situation. Besides the use of directional microphone arrays (e.g. Soede et al., 1993a,b) signal processing algorithms based on binaural-hearing strategies might be applied in hearing aids to obtain the necessary improvement of the signal-to-noise ratio (Kollmeier, 1992; Peissig, 1992; Bodden, 1993).

References

Bilsen, F.A. and Raatgever, J. (1973) Spectral dominance in 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.

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

Bodden, M. (1993) Modeling human sound-source localization and the cocktail-party-effect. Acta Acustica 1, 43-55.

Bronkhorst, A. (1993) Personal communication.

Colburn, H.S. and Durlach, N.I. (1978) Models of binaural interaction. In: E.C. Carterette and M.P. Friedman (Eds.), Handbook of perception , Vol. IV, Hearing, Academic, New York.

Genuit, K. (1991) Gehörgerechte Lärmbewertung. Fortschritte der Akustik, DAGA '91, 75-92.

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

Hellbrück, J. and Kilcher, H. (1993) Effects on mental tasks induced by noise recorded and presented via an artificial head system. Proc. 6th "Noise and Man", Nice, France.

Houtgast, T. and Aoki, S. (1994) Stimulus-onset dominance in the perception of binaural information. Hearing Research 72, 29-36.

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

Kollmeier, B. (1992) Noise reduction algorithms for hearing aids. Abstract 1st Europ. Conf. Audiol. , Cambridge, England.

Lindemann, W. (1986) Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization of stationary signals. J. Acoust. Soc. Am. 80, 1608-1622.

Notbohm, G., Schwarze, S. and Jansen, G. (1992) Noise evaluation based on binaural hearing. Proc. 14th Int. Congr. Acoust., Beijing, China, paper H2-2.

Peissig, J. (1992) Binaurale Hörgerätestrategien in komplexen Störschallsituationen. Doctoral dissertation, Göttingen, Germany.

Potter, J. (1993) On the binaural modelling of spaciousness in room acoustics. Doctoral dissertation, Delft Univ. Technol., Delft, The Netherlands.

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 Press, 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.

Raatgever, J. and van Keulen, W. (1992) Binaural time processing at high frequencies: the central spectrum model extended. Proc. 14th Int. Congr. Acoust., Beijing, paper H6-3.

Salomons, A.M., Bilsen, F.A. and Raatgever, J. (1991) Coloration of diotic and dichotic signals with multiple reflections. Proc. DAGA '91, Bochum, 481-484.

Soede, W., Berkhout, A.J. and Bilsen, F.A. (1993) Development of a directional hearing instrument based on array technology. J. Acoust. Soc. Am. 94, 785-798.

Soede, W., Bilsen, F.A. and Berkhout, A.J. (1993) Assessment of a directional microphone array for hearing-impaired listeners. J. Acoust. Soc. Am. 94, 799-808.

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.

Wightman, F.L. and Kistler, D.J. (1992) The dominant role of low-frequency interaural time differences in sound localization. J. Acoust. Soc. Am. 91, 1648-1661.