Hearing Research 256 (2009) 39–57

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Electrophysiological and psychophysical asymmetries in sensitivity to interaural correlation steps

Helge Lüddemann *, Helmut Riedel, Birger Kollmeier

Medizinische Physik, Carl von Ossietzky Universität Oldenburg, D-26111 Oldenburg, Germany

ARTICLE INFO

Article history:
Received 27 November 2008
Received in revised form 12 June 2009
Accepted 15 June 2009
Available online 23 June 2009

Keywords:
Binaural psychoacoustics
Spatial diffuseness
Scale transform
Late auditory evoked potentials
Sustained potential
Auditory scene analysis

ABSTRACT

The binaural auditory system's sensitivity to changes in the interaural cross correlation (IAC), as an indicator for the perceived spatial diffuseness of a sound, is of major importance for the ability to distinguish concurrent sound sources. In this article, we present electroencephalographical and corresponding psychophysical experiments with stepwise transitions of the IAC in continuously running noise.

Both the transient and sustained brain response, display electrophysiological correlates of specific binaural processing in humans. The transient late auditory evoked potentials (LAEP) systematically depend on the size of the IAC transition, the reference correlation preceding the transition, the direction of the transition and on unspecific context information from the stimulus sequence.

The psychophysical and electrophysiological data are characterized by two asymmetries. (1) Major asymmetry: for reference correlations of $+1$ and -1, psychoacoustical thresholds are comparatively lower, and the peak-to-peak-amplitudes of LAEP are larger than for a reference correlation of zero. (2) Minor asymmetry: for IAC transitions in the positive parameter range, perceptual thresholds are slightly better and peak-to-peak amplitudes are larger than in the negative range.

In all experimental conditions, LAEP amplitudes are linearly related to the dB scaled power ratio of correlated (N_0) versus anticorrelated (N_π) signal components. The voltage gain of LAEP per dB(N_0/N_π) closely corresponds to a constant perceptual distance between two correlations. We therefore suggest that activity in the auditory cortex and perceptual IAC sensitivity are better represented by the dB-scaled N_0/N_π power ratio than by the normalized IAC itself.

1. Introduction

In realistic acoustical environments, several concurrent sound sources are often present at the same time, each having a different spatial extent and being masked by diffuse ambient noise or reverberation. Binaural listening is generally thought to facilitate the ability to distinguish single sound sources of particular interest from others by their spatial position (Colburn, 1995; Bronkhorst, 2000; Faller and Merimaa, 2004; Beutelmann and Brand, 2006; Nix and Hohmann, 2007).

According to the duplex theory of sound, the interaural time difference (ITD) and the interaural level difference (ILD) between both ears are the most valuable physical signal parameters for human listeners to localize a single sound source in the azimuthal plane (Rayleigh, 1907; Yost and Gourevitch, 1987). In acoustically complex situations, however, the interaural timing and level disparities of the incident sound waves no longer provide consistent spatial information across frequency and time due to the superposition of sounds from different locations. In particular, in the presence of diffuse ambient noise ITDs and ILDs might become inconsistent even within small spectrotemporal portions of the signal and appear to be random variables as the noise level increases (Nix and Hohmann, 2006, 2007). As a consequence, the precise localization of distinct sound sources is substantially impaired (Saberi et al., 1998). Perceptually, this results in a broadening of the sound object's width in auditory space, i.e., the object is perceived as more or less diffuse.

Rather than investigating the effect of spatial diffuseness on performance in lateralization tasks (Saberi et al., 1998; Trahiotis et al., 2001), the aim of this contribution is to understand how the perceived diffuseness per se – or, as its related physical quantity, the interaural cross correlation (IAC) – is represented in the

---

Abbreviations: AFC, alternative forced choice; ASSR, auditory steady state response; BOLD, blood oxygen level dependent; EEG, electroencephalography; ERB, equivalent rectangular bandwidth (of an auditory filter); fMRI, functional magnetic resonance imaging; IAC, interaural cross correlation; ITD, interaural time difference; ILD, interaural level difference; JND, just noticeable difference; JNT, just noticeable transition; LAEP, late auditory evoked potential; MEG, magnetoencephalography; N1, negative deflection in the LAEP at about 100–130 ms after the stimulus; P2, positive deflection in the LAEP at about 200–230 ms after the stimulus; SDT, signal detection theory; SNR, signal-to-noise ratio; SP, sustained potential

* Corresponding author. Tel.: +49 441 798 5472; fax: +49 441 798 3902.
E-mail address: h.lueddemann@uni-oldenburg.de (H. Lüddemann).

0378-5955/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2009.06.010

auditory system. This question is assessed by a combination of electroencephalographical recordings (EEG) of late auditory evoked potentials (LAEP) and by closely corresponding psychophysical experiments, using stimuli with stepwise transitions of the IAC in continuously running broadband noise.

1.1. The normalized cross correlation and related quantities

The IAC between the signals $l(t)$ at the left and $r(t)$ at the right ear is defined as the normalized scalar product of l and r and hereby denoted by ρ,

$$\rho = \frac{\int l(t) \cdot r(t)\,dt}{\sqrt{\int l^2(t)\,dt} \cdot \sqrt{\int r^2(t)\,dt}}. \tag{1}$$

Due to normalization the range of ρ is restricted to the interval $[-1;+1]$. When presented via headphones, correlated noise signals ($\rho = +1$, also termed N_0) are perceived as a compact sound source at a central position in the head, whereas uncorrelated signals ($\rho = 0$, also termed N_u) are perceived as diffuse, i.e., they are associated with a continuum of simultaneously active sources between both ears. Anticorrelated signals ($\rho = -1$, also termed N_π) are often associated with two sound sources, one at the left and the other at the right ear. Hence, by setting ρ to an intermediate value, the overall spatial diffuseness of dichotic noise stimuli can be adjusted continuously, simulating the amount of either consistent or inconsistent information about the spatial distribution of sound sources in a complex acoustical environment. This allows one to systematically investigate how the binaural system deals with different degrees of diffuseness.

Noise stimuli with any desired ρ can be generated by mixing two orthonormal noise sources $a(t)$ and $b(t)$ in an appropriate ratio (Culling et al., 2001):

$$\begin{pmatrix} l \\ r \end{pmatrix} = \begin{pmatrix} a \\ \rho \cdot a + \sqrt{1-\rho^2} \cdot b \end{pmatrix} \tag{2}$$

Based on a descriptive analysis of binaurally masked tone detection thresholds, van der Heijden and Trahiotis (1997) suggested that the amount of masking produced by a noise having an arbitrary interaural correlation is equal to the addition of the masking effects produced by the diotic (N_0) and anticorrelated (N_π) constituents that compose the masker. Using this concept, they could successfully explain the dependence of binaural masking level differences (BMLD) on the interaural correlation for various masker bandwidths.

Motivated by this additivity of masking, one can also use an alternative mixing formula in which stimuli are obtained as a mixture of a diotic noise (N_0) and an antiphasic noise (N_π), which are again built from two orthonormal noise sources $a(t)$ and $b(t)$:

$$\begin{pmatrix} l \\ r \end{pmatrix} = \sqrt{\frac{1+\rho}{2}} \begin{pmatrix} a \\ a \end{pmatrix} + \sqrt{\frac{1-\rho}{2}} \begin{pmatrix} b \\ -b \end{pmatrix} \tag{3}$$

or:

$$N_\rho = \sqrt{\frac{1+\rho}{2}}\, N_0 + \sqrt{\frac{1-\rho}{2}}\, N_\pi \tag{4}$$

The mixing formula (4) illustrates that the IAC – and thus the diffuseness – of any dichotic signal N_ρ is entirely determined by the ratio of its correlated (N_0) versus its anticorrelated (N_π) components. This ratio of mixing coefficients provides an alternative representation of IAC,

$$\tilde{\rho} = 10 \cdot \log \frac{1+\rho}{1-\rho}. \tag{5}$$

Except for a proportionality factor of 10, $\tilde{\rho}$ is the Fisher Z-transform of ρ. $\tilde{\rho}$ is a monotonic nonlinear function of the normalized IAC ρ. It is identical to Durlach's "equivalent signal-to-noise ratio" (equivalent SNR) in an $N_\pi S_0$ BMLD paradigm. However, in order to prevent confusion, we denote $\tilde{\rho}$ as the "dB(N_0/N_π) scaled IAC" because Durlach et al. (1986) introduced several different formulae for the computation of an equivalent SNR any of which corresponds to a particular interaural configuration of masker and signal phase in other BMLD experiments, e.g., $N_u S_0$ or $N_0 S_m$.

1.2. Psychoacoustics on interaural cross correlation

The just noticeable difference between two values of normalized IAC (ρ-JND) critically depends on the correlation of the reference stimulus, ρ_{ref}. For stimulus bandwidths greater than 1 ERB (Moore et al., 1988), ρ-JNDs at $\rho_{\text{ref}} = +1$ are between 0.02 and 0.057, while ρ-JNDs at $\rho_{\text{ref}} = 0$ range from 0.3 up to 0.72 (Gabriel and Colburn, 1981; Koehnke et al., 1986; Akeroyd and Summerfield, 1999; Culling et al., 2001; Boehnke et al., 2002), i.e., psychoacoustical discrimination thresholds are about 10 times lower for $\rho_{\text{ref}} = +1$ than for $\rho_{\text{ref}} = 0$. For intermediate ρ_{ref}, there is a nonlinear decrease of the ρ-JND as ρ_{ref} increases from 0 to +1 (Pollack and Trittipoe, 1959a; Culling et al., 2001, 2003). The differences between the absolute JND values that have been reported in the literature are presumably due to different spectral stimulus properties and experimental techniques (Gabriel and Colburn, 1981; Akeroyd and Summerfield, 1999; Culling et al., 2001).

Measuring the ρ-JND for $\rho_{\text{ref}} = -1$, Boehnke et al. (2002) found that thresholds were markedly lower than for $\rho_{\text{ref}} = 0$ but were twice as large as for $\rho_{\text{ref}} = +1$. In addition, for $\rho_{\text{ref}} = 0$ they reported lower thresholds for positive than for negative deviant correlations ρ_{dev}. Also the cumulative d'-functions by Culling et al. (2003) indicate that the discriminability of two stimuli with different correlations is generally worse in the negative than in the positive range of ρ.

In the above-mentioned IAC discrimination experiments listeners had to rely on auditory memory in order to distinguish stimuli with static but different correlations ρ_{ref} and ρ_{dev} which were separated in time. In realistic acoustical situations, however, there are more or less rapid IAC transitions within the ongoing signal which can serve as an additional dynamic cue for binaural scene analysis. In analogy to the importance of common onsets in the monaural case, such rapid changes of the interaural signal properties might provide even more salient cues for binaural object separation than the memory-based comparison of signals with silence in between, or the comparison of quasi-static interaural parameters in temporally subsequent segments of an ongoing sound.

Nevertheless, the detection of stepwise IAC transitions and the discrimination of stimuli with static IAC seem to be of quite similar character: Dajani and Picton (2006) periodically switched the IAC between zero and a positive deviant correlation $\Delta\rho$ at a switch rate of 4 Hz, i.e., with 8 IAC transitions per second. They reported that such stimuli with rectangular modulations of the IAC could be distinguished from an uncorrelated noise if $\Delta\rho$ exceeded 0.31. Grantham (1982) investigated the detectability of sinusoidal IAC modulations at various modulation rates. For a modulation rate of 5 Hz, his results indicated that discrimination of the modulated stimuli against uncorrelated reference noise was possible at a modulation amplitude of about 0.45, i.e., with the IAC oscillating between −0.45 and +0.45.

In experiments with binaural gaps, i.e., brief segments of deviant correlation ρ_{dev} which were temporally flanked by reference segments with an IAC of ρ_{ref}, Boehnke et al. (2002) found qualitatively the same dependence on ρ_{ref} for the gap duration thresholds as for the ρ-JNDs in corresponding discrimination tasks with static IAC.

In summary, psychoacoustical literature data suggest that IAC sensitivity can be characterized by two asymmetries: First, there