Behavioural Brain Research 292 (2015) 157–166

Contents lists available at ScienceDirect

Behavioural Brain Research

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

Electrophysiological correlates of cocktail-party listening

Jörg Lewald[a,b,*], Stephan Getzmann[a,b]

[a] Auditory Cognitive Neuroscience Laboratory, Department of Cognitive Psychology, Ruhr University Bochum, D-44780 Bochum, Germany
[b] Leibniz Research Centre for Working Environment and Human Factors, Ardeystraße 67, D-44139 Dortmund, Germany

HIGHLIGHTS

- Sound localization in the presence of multiple sources was studied using ERPs.
- Multiple sources engage a different pattern of neural activity than single sources.
- A strong N2 component occurs with multiple but not single sources.
- At N2, multiple sources modulate activation in dorso–frontal and cingulate cortices.
- An anterior contralateral N2ac component reflects focusing of spatial attention.

ARTICLE INFO

Article history:
Received 16 April 2015
Received in revised form 12 June 2015
Accepted 14 June 2015
Available online 16 June 2015

Keywords:
Sound localization
Auditory scene analysis
Cocktail party effect
Selective spatial attention
Auditory evoked potentials

ABSTRACT

Detecting, localizing, and selectively attending to a particular sound source of interest in complex auditory scenes composed of multiple competing sources is a remarkable capacity of the human auditory system. The neural basis of this so-called "cocktail-party effect" has remained largely unknown. Here, we studied the cortical network engaged in solving the "cocktail-party" problem, using event-related potentials (ERPs) in combination with two tasks demanding horizontal localization of a naturalistic target sound presented either in silence or in the presence of multiple competing sound sources. Presentation of multiple sound sources, as compared to single sources, induced an increased P1 amplitude, a reduction in N1, and a strong N2 component, resulting in a pronounced negativity in the ERP difference waveform (N2d) around 260 ms after stimulus onset. About 100 ms later, the anterior contralateral N2 subcomponent (N2ac) occurred in the multiple-sources condition, as computed from the amplitude difference for targets in the left minus right hemispaces. Cortical source analyses of the ERP modulation, resulting from the contrast of multiple vs. single sources, generally revealed an initial enhancement of electrical activity in right temporo–parietal areas, including auditory cortex, by multiple sources (at P1) that is followed by a reduction, with the primary sources shifting from right inferior parietal lobule (at N1) to left dorso–frontal cortex (at N2d). Thus, cocktail-party listening, as compared to single-source localization, appears to be based on a complex chronology of successive electrical activities within a specific cortical network involved in spatial hearing in complex situations.

1. Introduction

One of the most astonishing perceptual capacities of humans as well as the vast majority of hearing animals is their ability to detect, localize, and selectively attend to a particular sound source of interest in complex acoustic environments. This ability has been termed "cocktail-party effect" [1]. Despite its essential importance for both communication and orientation in space, the neural basis of this phenomenon has remained long unknown, while the mechanisms involved in the localization of an isolated sound source in an otherwise silent and anechoic environment have been investigated in detail (for review, see, e.g., [2,3]). However, particularly over the last decade, the specific features of spatial hearing under more naturalistic "cocktail-party" conditions have increasingly become the subject of neuroscience research, not only in human subjects [4,5], but also in non-mammalian animals, such as birds [6], frogs [7], and insects [8].

In human subjects, the results of many studies, using various imaging methods in combination with tasks of localization of isolated sound sources (for review, see [9,10]), have clearly indicated the involvement of the postero–dorsal auditory processing stream, comprising posterior parts of auditory cortex, posterior superior temporal gyrus, inferior parietal lobule, and superior frontal sulcus, in the successful solution of these tasks [2]. Beyond that, there

* Corresponding author. Fax: +49 2343205137.
E-mail address: joerg.lewald@rub.de (J. Lewald).

http://dx.doi.org/10.1016/j.bbr.2015.06.025

is also evidence that the components of the antero–ventral auditory stream, namely the anterior temporal lobe and the inferior frontal lobe, are activated during tasks of sound localization [11,12]. It has been hypothesized that this antero–ventral activation may reflect the shared analysis of object-feature information and spectral localization cues [11,12].

Compared with the localization of a single sound source in silence, the “cocktail-party” situation is substantially more demanding for the auditory system. The extraction and localization of a stimulus of interest among multiple other sounds require, in addition to the genuine spatial analysis, the simultaneous analyses of several other acoustic features, such as pitch and timbre. Moreover, increased attentional effort is necessary to focus on the stimulus of interest while inhibiting concurrent sound information. Thus, one may assume that the neural network involved in solving the “cocktail-party” task differs, at least to some extent, from that involved in single-source localization. In an initial approach to this issue, Zündorf et al. [4] used functional magnetic resonance imaging (fMRI) in combination with presentation of single and multiple virtual sound sources. The results demonstrated crucial roles of bilateral posterior superior temporal gyrus, left anterior insula, supplementary motor area, and fronto–parietal network in auditory stream segregation, of the left planum temporale in extracting the sound of interest among acoustical distracters, and of the right precuneus in orienting spatial attention to the target sound. In a subsequent study using voxel-based lesion-behaviour mapping (VLBM) analyses in stroke patients, Zündorf et al. [5] compared the patients’ abilities to localize acoustic targets in a single-source and in a multi-source free-field setup to uncover the brain areas associated with a deficit in localization in the presence of multiple distracter sound sources rather than localization of individually presented sound sources. This study revealed a fundamental role of the right planum temporale in the “cocktail-party” task, as well as some involvement of left inferior frontal and pre- and postcentral areas. These two appoaches [4,5] provided initial evidence that, beyond the known network for localization of single sources in otherwise silent surroundings, specific networks are particularly involved in the solution of the “cocktail-party” problem, namely areas subserving spectro–temporal analyses for the effective segregation of multiple sound streams from various locations.

While the previous studies [4,5] showed the general importance of specific brain areas for sound localization in a “cocktail- party” situation, the present study aimed to clarify the exact time course of these processes, which might involve distinct networks at different points in time (cf. [12]). For this purpose, we recorded event-related potentials (ERPs) and employed electrical neuroimaging (electrotomography) with high temporal resolution in the millisecond range and low, but acceptable, spatial resolution in the range of several millimeters [13–15]. As in our previous studies [4,5], localization of a target sound in the presence of multiple distractor sources was contrasted with single-source localization to reveal the ERPs associated with and cortical areas involved in the extraction of the location of the acoustic event of interest from other distracting sound sources within an auditory scene.

As an additional approach, we addressed the role of auditory attention for sound localization in a “cocktail-party” situation, namely the process of directing the focus of spatial attention toward the stimulus of interest. As a potential correlate of auditory spatial attention, we investigated the so-called N2ac component that has been recently proposed to be an auditory analogue of the N2pc [16]. The N2pc is a posterior contralateral subcomponent of the visual N2 that is associated with lateralized shifts of attention toward a target surrounded by multiple distractors [17–21]. In their recent study, Gamble and Luck [16] found an anterior contralateral negativity in the N2 latency range, the N2ac, that was elicited when subjects indicated whether one of two sounds presented simultaneously to the left and right was a pre-defined target. In the present study, we extended the approach of Gamble and Luck [16], and tested the localization of one target source in the presence of three distractor sources, presented in left (−30°; −90°) and right (30°; 90°) hemispaces. By contrasting the ERP for targets in left and right hemispaces, we explored the existence of the N2ac component that should become manifest as anterior contralateral negativity in the N2 latency range.

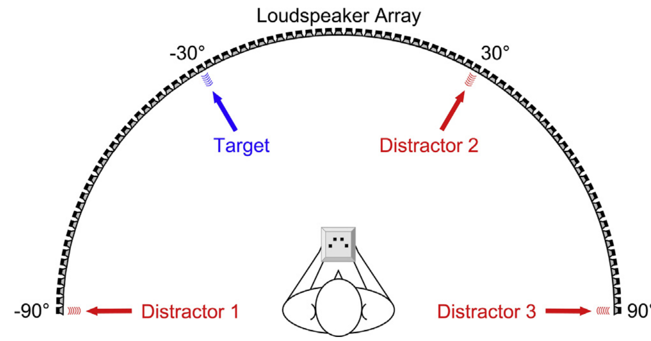

Fig. 1. Experimental set-up. Target sound sources were either presented in isolation (single-source condition) or simultaneously with three other sound sources (multiple-sources condition) from four loudspeakers located at −90° and −30° to the left and 30° and 90° to the right (the total array, covering 180°, was composed of 91 loudspeakers). Subjects had to localize the position of the target by pressing one out of four response buttons on a response box.

2. Methods

2.1. Subjects

Fourty healthy subjects (20 men, 20 women; mean age 24.0, SE 0.5, range 19–34 years) participated in the study. All subjects wrote with their right hand. Standard pure-tone audiometry (Oscilla USB100, Inmedico, Lystrup, Denmark) was obtained from all subjects (11 frequencies from 0.125 to 8 kHz presented to either ear). Thresholds of all subjects were normal (≤25 dB hearing level) for each frequency tested. The study conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the British Medical Journal (18 July 1964) and was approved by the Ethical Committee of the Leibniz Research Centre for Working Environment and Human Factors, Dortmund. All subjects gave their written informed consent to participate in the study.

2.2. Apparatus, stimuli and procedure

Experiments were conducted in a dimly illuminated, anechoic and sound-proof room (for details, see [22]). Subjects sat on a comfortable chair. The position of the subject’s head was held constant by a chin rest. In front of the subject, a semicircular array of 91 broad-band loudspeakers (SC 5.9, Visaton, Haan, Germany) was mounted at a distance of 1.5 m from the centre of the subject’s head at ear level, thus covering 180° in steps of 2° (for details, see [23]). Only four loudspeakers, located at azimuthal positions of −90°, −30° (to the left), 30°, and 90° (to the right of the subject’s median plane) were used in the present study (Fig. 1). A red light-emitting diode (LED; diameter 3 mm, luminance 0.025 mcd) located in the subject’s median plane at eye level served as a fixation target and was always on.

Four animal vocalizations were used as auditory stimuli. The original sound files (‘birds chirping’; ‘dog barking’; ‘frog’; ‘sheep’) were selected from an online sound library [24], based on their familiarity and recognizability. Sounds were adjusted to a constant duration of 600 ms using the software Cool Edit 2000 (Syntrillium Software Corporation, Phoenix, AZ, USA). As the original sound files