BRAIN RESEARCH 1387 (2011) 116–124

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Neural generators underlying concurrent sound segregation

Stephen R. Arnott[a,*], Tim Bardouille[b], Bernhard Ross[a,c,d], Claude Alain[a,e]

[a]Rotman Research Institute, Baycrest Centre, Toronto, Ontario, Canada M6A 2E1
[b]Institute for Biodiagnostics (Atlantic), Halifax, Nova Scotia, Canada B3H 3A7
[c]Institute of Medical Sciences, University of Toronto, Ontario, Canada M8V 2S4
[d]Department for Medical Biophysics, University of Toronto, Ontario, Canada
[e]Department of Psychology, University of Toronto, Ontario, Canada M8V 2S4

ARTICLE INFO

Article history:
Accepted 19 February 2011
Available online 26 February 2011

Keywords:
Concurrent sound segregation
ORN
MEG
Multiple auditory objects
Evoked field

ABSTRACT

Although an object-based account of auditory attention has become an increasingly popular model for understanding how temporally overlapping sounds are segregated, relatively little is known about the cortical circuit that supports such ability. In the present study, we applied a beamformer spatial filter to magnetoencephalography (MEG) data recorded during an auditory paradigm that used inharmonicity to promote the formation of multiple auditory objects. Using this unconstrained, data-driven approach, the evoked field component linked with the perception of multiple auditory objects (i.e., the object-related negativity; ORNm), was found to be associated with bilateral auditory cortex sources that were distinct from those coinciding with the P1m, N1m, and P2m responses elicited by sound onset. The right hemispheric ORNm source in particular was consistently positioned anterior to the other sources across two experiments. These findings are consistent with earlier proposals of multiple auditory object detection being associated with generators in the auditory cortex and further suggest that these neural populations are distinct from the long latency evoked responses reflecting the detection of sound onset.

1. Introduction

A major goal of hearing research is to understand how listeners extract a particular event or sequence amid the temporally-overlapping mixture of acoustic information arriving at their ears (i.e., auditory scene analysis; Alain and Bernstein, 2008; Bregman, 1990). In recent years, the matter of auditory scene analysis has taken on renewed interest with advances in neuroimaging technologies and the desire to understand aged and diseased auditory systems.

One account of auditory scene analysis that has grown in popularity over the past decade is the object-based model (Alain and Arnott, 2000; Griffiths and Warren, 2004; Scholl, 2001; Shinn-Cunningham, 2008). Object-based models generally hold that auditory attention is governed by gestalt-like grouping principles (Koffka, 1935) that enable a listener to (actively or passively) group the auditory features that surround them. Although much of the object-based attention literature is largely theoretical, often drawing upon analogies from vision (Scholl, 2001; Shinn-Cunningham, 2008), attempts

* *Corresponding author*. Fax: + 1 416 785 2862.
E-mail address: sarnott@rotman-baycrest.on.ca (S.R. Arnott).

0006-8993/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2011.02.062

have been made to find objective measures of auditory object formation (e.g., Alain and Woods, 1997; Alain et al., 2001; Arnott and Alain, 2002; Bidet-Caulet et al., 2007).

A stimulus that has been particularly amenable to auditory object research is the mistuned harmonic sound (Moore et al., 1985; Moore et al., 1986). Typically, a complex tone comprised of integer harmonics of its fundamental frequency (f_0) is perceived as a single sound. However, when one of the lower harmonics is mistuned by a certain degree (usually 10–16% of its original frequency), listeners report hearing *two* sounds, a pure tone "standing out" from a complex tone background. With increasing amounts of mistuning, listeners are more likely to report hearing two sounds. Because the phenomenon involves temporally synchronous sounds with very similar acoustical structures, the mistuned harmonic paradigm is well suited for studying the neural correlates of concurrent sound segregation since the perception of multiple auditory objects can be dissociated from most sensory components simply by examining the difference between the tuned and mistuned stimuli. In fact, when the evoked responses associated with a non-mistuned harmonic complex are subtracted from those associated with a mistuned one, a neural marker known as the object-related negativity (ORN) is revealed as a negative deflection 160–200 ms after sound onset (Alain et al., 2001). The ORN (and its magnetic counterpart, the ORNm) can be elicited whether listeners actively or passively listen to the sounds (Alain et al., 2001), is found in both young and old listeners (Alain and McDonald, 2007), and seems to be little affected by task difficulty or visual attention load (Dyson et al., 2005).

Attempts to localize the neural generator(s) of the ORN using equivalent current dipole (ECD) source have suggested that the ORN is generated in or near the auditory cortex (Alain et al., 2001), consistent with that area's role in auditory stream segregation and object perception (Alain et al., 2005; Alain and Bernstein, 2008; Bee and Klump, 2004; Deike et al., 2010; Itatani and Klump, 2009; Micheyl et al., 2005; Schadwinkel and Gutschalk, 2010; Snyder et al., 2006; Sussman et al., 1999). Exactly what aspect of object perception the ORN is indexing remains an open question, however a reasonable hypothesis is that it relates to the detection of the second object's presence (i.e., its onset). Given that the mistuned and tuned harmonics of the ORN stimulus onset at the same time, one might expect that the detection of the second "object" would be somewhat delayed (relative to the detection of the overall stimulus) since a certain degree of sound sampling would be needed before the inharmonicity could be realized. It is well known in the field of event-related potentials that the onset of a sound is associated with an obligatory "P1–N1–P2" long-latency response beginning around 50 ms after the sound has been perceived (Knight et al., 1980; Näätänen and Picton, 1987). Accordingly, it is possible that the ORN component may itself be a manifestation of a temporally delayed P1, N1, and/or P2 in response to the detection of the mistuned harmonic (Alain et al., 2001; Alain and McDonald, 2007; Alain et al., 2009). The plausibility of this "delayed onset" hypothesis is bolstered by dipole source localization studies that place the neural generators of these P1, N1 and P2 responses in regions (Hari et al., 1987) proximal to where the ORN is believed to occur (i.e., the Sylvian fissure). However, due to the limitations of ECD modeling, including its reliance on user imposed constraints (see Fender, 1987; Scherg and Von Cramon, 1985), it remains unclear whether ORN generators really are independent from N1 generators.

In the present study, we sought to localize generators of the ORN as well as the obligatory onset response by applying event-related synthetic aperture magnetometry (ER-SAM; Cheyne et al., 2006; Robinson and Vrba, 2004) to two mistuned harmonic MEG data sets (Alain and McDonald, 2007; Alain et al., 2009). Unlike ECD source modeling, the ER-SAM algorithm does not require *a priori* assumptions about the number or location of sources (e.g., basing the ORN model on an N1 source). Rather, ER-SAM uses a minimum variance beamformer algorithm as a spatial filter to estimate neuronal activity from any location in the brain. Compared to other analytic techniques such as the minimum norm estimate [MNE] approach (Hämäläinen and Ilmoniemi, 1994), beamforming employs two innovations that potentially result in increased spatial resolution. First, in addition to physical sensitivity properties of the sensor, signal properties (i.e., temporal covariances between magnetic field signals) are incorporated in the algorithm to increase the spatial resolution (Van Veen et al., 1997). Second, normalization based on a noise estimate results in almost uniform sensitivity across the brain volume, allowing for imaging of deep source activity (Vrba and Robinson, 2001).

Using information from all MEG sensors, the SAM approach involves dividing the entire brain volume into a grid of nodes, and then using the beamformer to enhance the signal at each node while also suppressing the signals from the other nodes (Huang et al., 2004). The time course of current source activity at each node/voxel of the brain is then estimated as a weighted linear combination of the magnetic field measured at all MEG sensor positions and plotted as a pseudo-Z value. Voxels showing maximum pseudo-Z values can then be selected as representatives of the center of gravity of neural sources. Unlike in functional magnetic resonance imaging where the spatial distribution of activation maps reflects (in first order approximation) the extent of cortical activation, the spatial distribution of the ER-SAM MEG map relates more to limited spatial resolution. To date, ER-SAM analyses have been successfully used to analyze mid and long latency auditory evoked transient responses (Du et al., in press; Steinsträter et al., 2007).

If the ORN is in fact indexing a delayed onset response, then one might expect ER-SAM analyses to provide ORNm source locations that are the same as those attained for P1m, N1m or P2m sources, but simply delayed in time. Alternatively, the ORN may be indexing something different from what those responses reflect, thereby activating different neural generators. For example, auditory object analysis is also known to invoke activity in brain regions anterior to the primary auditory cortex during sound categorization tasks (Alain et al., 2001; Alain and Bernstein, 2008; Arnott et al., 2004; Tian et al., 2001), with more anterior regions being activated as sounds become increasingly complex (Scott, 2005). Similarly, brain regions posterior to the auditory cortex may play a role in auditory object formation by participating in complex acoustic analysis required for sound segmentation (Harms et al., 2005), disambiguating co-occurring environmental sounds that originate from the same region of space (Zatorre