NeuroImage: Clinical 7 (2015) 699–708

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Functional neuroanatomy of auditory scene analysis in Alzheimer's disease

Hannah L. Golden [a], Jennifer L. Agustus [a], Johanna C. Goll [a], Laura E. Downey [a], Catherine J. Mummery [a], Jonathan M. Schott [a], Sebastian J. Crutch [a], Jason D. Warren [a,*]

[a] Dementia Research Centre, UCL Institute of Neurology, University College London, London, UK

ARTICLE INFO

Article history:
Received 20 October 2014
Received in revised form 16 January 2015
Accepted 24 February 2015
Available online 28 February 2015

Keywords:
Auditory scene analysis
Cocktail party effect
Alzheimer's disease
fMRI

ABSTRACT

Auditory scene analysis is a demanding computational process that is performed automatically and efficiently by the healthy brain but vulnerable to the neurodegenerative pathology of Alzheimer's disease. Here we assessed the functional neuroanatomy of auditory scene analysis in Alzheimer's disease using the well-known 'cocktail party effect' as a model paradigm whereby stored templates for auditory objects (e.g., hearing one's spoken name) are used to segregate auditory 'foreground' and 'background'. Patients with typical amnestic Alzheimer's disease ($n = 13$) and age-matched healthy individuals ($n = 17$) underwent functional 3T-MRI using a sparse acquisition protocol with passive listening to auditory stimulus conditions comprising the participant's own name interleaved with or superimposed on multi-talker babble, and spectrally rotated (unrecognisable) analogues of these conditions. Name identification (conditions containing the participant's own name contrasted with spectrally rotated analogues) produced extensive bilateral activation involving superior temporal cortex in both the AD and healthy control groups, with no significant differences between groups. Auditory object segregation (conditions with interleaved name sounds contrasted with superimposed name sounds) produced activation of right posterior superior temporal cortex in both groups, again with no differences between groups. However, the cocktail party effect (interaction of own name identification with auditory object segregation processing) produced activation of right supramarginal gyrus in the AD group that was significantly enhanced compared with the healthy control group. The findings delineate an altered functional neuroanatomical profile of auditory scene analysis in Alzheimer's disease that may constitute a novel computational signature of this neurodegenerative pathology.

1. Introduction

Decoding the auditory world poses a formidable problem of neural computation. Our brains normally solve this problem efficiently and automatically but the neural basis of 'auditory scene analysis' remains incompletely understood. The disambiguation of sound sources within the complex mixture that generally arrives at our ears is an essential prerequisite for identification of those sources and a fundamental task of auditory scene analysis (Bregman, 1994). One of the best known instances of this process in action is the so-called 'cocktail party effect' whereby our own name spoken across a noisy room captures attention and may even lead to successful tracking of the relevant conversation against the surrounding babble (Cherry, 1953; Moray, 1959). The cocktail party effect is a celebrated example of a much wider category of auditory phenomena that depend on generic computational processes that together segregate an acoustic target or 'foreground' sound from the acoustic 'background': these processes are likely to include representation of spectral and temporal regularities in the sound mixture and matching to previously stored auditory 'templates' (for example, specific speech or vocal sounds) prior to engagement of attentional resources (Billig et al., 2013; Griffiths and Warren, 2002; Kumar et al., 2007). Functional neuroimaging studies to define neuroanatomical substrates of auditory scene analysis in the healthy brain have implicated a distributed, dorsally directed cortical network including planum temporale and posterior superior temporal gyrus, supramarginal gyrus, intraparietal sulcus and prefrontal projection targets (Dykstra et al., 2011; Gutschalk et al., 2007; Hill and Miller, 2010; Kondo and Kashino, 2009; Kong et al., 2014; Linden et al., 1999; Overath et al., 2010; Wilson et al., 2007; Wong et al., 2009). While frontal cortex is thought to drive top-down attentional processes (Hill and Miller, 2010; Obleser et al., 2007; Schönwiesner et al., 2007), the precise role of parietal cortex in auditory scene analysis is more contentious and might include primary labelling of salient events (Cohen, 2009; Downar et al., 2000), integration of signal representations for programming behavioural responses (Cusack, 2005; Lee et al., 2014) or

* Corresponding author. Tel.: +44 203 448 4773; fax: +44 203 448 3104.
E-mail address: jason.warren@ucl.ac.uk (J.D. Warren).

http://dx.doi.org/10.1016/j.nicl.2015.02.019

attentional modulation (Hill and Miller, 2010; Nakai et al., 2005). With particular reference to the cocktail party effect, speech intelligibility has been shown to engage more ventral and anterior superior temporal cortex in the dominant hemisphere (Scott et al., 2000), but is influenced by the nature of the background masker (speech versus non-speech: Scott and McGettigan, 2013; Scott et al., 2009). Lexical processes may modulate auditory scene analysis, perhaps via template matching algorithms (Billig et al., 2013; Griffiths and Warren, 2002) as well as additional parietal and prefrontal mechanisms engaging in speech in noise processing, particularly under conditions of increased attentional demand (Binder et al., 2004; Davis et al., 2011; Nakai et al., 2005; Scott et al., 2004; Scott and McGettigan, 2013).

On behavioural as well as neuroanatomical grounds, the computational processing required for auditory scene analysis is likely to be particularly vulnerable to the neurodegenerative disease process in Alzheimer's disease (AD). Patients with AD commonly experience difficulties in following conversations under degraded listening conditions such as a busy room or noisy telephone line. Both generic deficits of central auditory processing and specific deficits of auditory scene analysis have been demonstrated in AD (Gates et al., 1996, 2008, 2011; Golden et al., 2015; Goll et al., 2011, 2012; Golob et al., 2007, 2009; Kurylo et al., 1993; Strouse et al., 1995); these develop early in the course of disease and are likely to interact with impairments of attention and working memory (Conway et al., 2001; Goll et al., 2012; Stopford et al., 2012). Deficits of auditory scene analysis are in accord with the neuroanatomy of AD, which blights a large-scale, functionally coherent brain network linking mesial temporal lobe structures with retrosplenial, temporo-parietal and medial prefrontal cortices (Buckner et al., 2008; Greicius and Menon, 2004; Raichle et al., 2001; Seeley et al., 2009). Regional deposition of pathogenic proteins, hypometabolism and atrophy within this network in AD closely overlaps regions implicated in auditory scene analysis and speech-in-noise processing in the healthy brain, and involvement of temporo-parietal cortical junction zones is likely to be particularly pertinent (Herholz et al., 2002; Scahill et al., 2002; Warren et al., 2012). Indeed, modulation of activity in these areas has been linked to the efficiency of speech-in-noise processing even in apparently healthy older individuals (Wong et al., 2009). However, the pathophysiology of this culprit brain network in AD remains to be worked out in detail. While involvement of this network is relatively selective in AD, it is unlikely that the network behaves as an amorphous unit (Warren et al., 2012); moreover its core function or functions have not been defined. Although it has been designated the 'default mode network', showing correlated activity in the healthy 'resting' brain and deactivation with certain tasks (Buckner et al., 2008; Raichle et al., 2001; Shulman et al., 1997), this network has also been implicated in various 'active' processes including maintenance of internal sensory representations (Buckner et al., 2008; Buckner and Carroll, 2007; Spreng and Grady, 2010; Zvyagintsev et al., 2013) and more specifically in aspects of auditory scene analysis, both in the healthy brain (Salvi et al., 2002; Wong et al., 2009; Zündorf et al., 2013) and in patients with AD (Goll et al., 2012).

Here we used the cocktail party effect to delineate the functional neuroanatomy of auditory scene analysis in a cohort of patients with AD in relation to healthy older individuals. Previous work in AD has addressed psychophysical deficits of auditory scene analysis using relatively simple paradigms and structural neuroanatomical correlation (Gates et al., 2008, 2011; Goll et al., 2012). In this study we set out to use a realistic auditory scene analysis paradigm in the context of fMRI, in order to probe functional brain mechanisms directly. This paradigm was motivated by a cognitive model of the cocktail party effect according to which stored templates for auditory objects (e.g., spoken words) are used to disambiguate those objects from other sounds in the environment during parsing of the auditory scene (segregation of auditory 'foreground' and 'background': Griffiths and Warren, 2002). We used participant's own names as salient acoustic targets (Moray, 1959; Wood and Cowan, 1995) against naturalistic multi-talker babble; a sparse fMRI acquisition protocol to minimise confounding effects engendered by streaming auditory stimuli against scanner noise (Hall et al., 1999); and a passive-listening design to minimise any confounding effects from output task in these cognitively impaired patients. Based on previous neuroanatomical work in the healthy brain and in AD, we hypothesised that patients with AD and healthy older individuals would show similar profiles of auditory cortex activation by sound and representation of name identity per se; but that AD would have a distinct pathophysiological signature during auditory scene analysis, in temporo-parietal cortical regions separable from more anterior superior temporal cortex engaged by name identity coding (Dykstra et al., 2011; Goll et al., 2012; Overath et al., 2010; Scott et al., 2000, 2009; Wong et al., 2009). In particular, we hypothesised that AD would produce an altered interaction of auditory name template matching with object segregation underpinning the cocktail party effect.

2. Methods

2.1. Participants

Thirteen consecutive patients (mean (standard deviation) age 66 (5.8) years; five female) fulfilling consensus clinical criteria for early to moderately severe, typical Alzheimer's disease (AD) led by predominant episodic memory loss with additional cognitive dysfunction (Dubois et al., 2007) and 17 age-matched healthy individuals (68 (3.9) years; seven female) with no history of neurological or psychiatric illness participated in the study. All participants were right-handed and no participant had a clinical history of peripheral hearing loss; none was a professional musician. Detailed general neuropsychological assessment in the AD group corroborated the clinical diagnosis in all cases; demographic, clinical and neuropsychological details for the experimental groups are summarised in Table 1. At the time of participation, 12 patients were receiving symptomatic treatment with an acetylcholinesterase inhibitor (one was also receiving memantine). CSF examination was undertaken in six patients with AD and revealed a total tau: beta-amyloid ratio >1 (compatible with underlying AD pathology) in all cases. All participants gave informed consent in accordance with the Declaration of Helsinki.

2.2. Assessment of peripheral hearing

All participants had pure-tone audiometry using a procedure adapted from a commercial screening audiometry software package (AUDIO-CDTM®, http://www.digital-recordings.com/audiocd/audio.html). The test was administered via headphones from a notebook computer in a quiet room. Five frequency levels (500, 1000, 2000, 3000, 4000 Hz) were assessed: at each frequency, participants were presented with a continuous tone that slowly and linearly increased in intensity. Participants were instructed to indicate as soon as they were sure they could detect the tone; this response time was measured and stored for offline analysis. Hearing was assessed in the right ear in each participant.

2.3. Experimental design and stimuli

In designing the experimental paradigm we manipulated two key components of the cocktail party effect: separation of a particular 'foreground' auditory object (a spoken word) from a complex sound mixture or acoustic 'background'; and matching of foreground object (own name) identity with a previously stored 'template'. In order to isolate the neural processes involved in these computations, we created two closely matched auditory baseline conditions: by presenting 'foreground' sounds interleaved with (rather than superimposed on) the acoustic background; and by spectral rotation of participants' spoken names to generate acoustically similar but unfamiliar (and unintelligible) sound objects. Under this design, the cocktail party effect (detection of own