



You talkin' to me? Communicative talker gaze activates left-lateralized superior temporal cortex during perception of degraded speech

Carolyn McGettigan^{a,b,*}, Kyle Jasmin^b, Frank Eisner^{b,c}, Zarinah K. Agnew^{b,d}, Oliver J. Josephs^{b,e}, Andrew J. Calder^f, Rosemary Jessop^b, Rebecca P. Lawson^{b,e}, Mona Spielmann^b, Sophie K. Scott^b

^a Department of Psychology, Royal Holloway University of London, Egham Hill, Egham TW20 0EX, UK

^b Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

^c Donders Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, Netherlands

^d Department of Otolaryngology, University of California, San Francisco, 513 Parnassus Avenue, San Francisco, CA, USA

^e Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK

^f MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK

ARTICLE INFO

Keywords:

Gaze
FMRI
Speech comprehension
Laterality

ABSTRACT

Neuroimaging studies of speech perception have consistently indicated a left-hemisphere dominance in the temporal lobes' responses to intelligible auditory speech signals (McGettigan and Scott, 2012). However, there are important communicative cues that cannot be extracted from auditory signals alone, including the direction of the talker's gaze. Previous work has implicated the superior temporal cortices in processing gaze direction, with evidence for predominantly right-lateralized responses (Carlin & Calder, 2013). The aim of the current study was to investigate whether the lateralization of responses to talker gaze differs in an auditory communicative context. Participants in a functional MRI experiment watched and listened to videos of spoken sentences in which the auditory intelligibility and talker gaze direction were manipulated factorially. We observed a left-dominant temporal lobe sensitivity to the talker's gaze direction, in which the left anterior superior temporal sulcus/gyrus and temporal pole showed an enhanced response to direct gaze – further investigation revealed that this pattern of lateralization was modulated by auditory intelligibility. Our results suggest flexibility in the distribution of neural responses to social cues in the face within the context of a challenging speech perception task.

1. Introduction

Spoken communication can only be described in part by reference to the exchange of linguistic messages. Natural conversation often occurs face-to-face, where interlocutors display facial expressions, gestures and non-verbal vocalizations (such as laughter) in order to enhance understanding, and to signal social cues such as mood, affiliation and intent. However, to date, relatively little is known about how the brain processes social and linguistic cues within the same communicative context.

Neuroimaging studies of auditory speech intelligibility in the healthy adult brain have attempted to isolate the neural responses to intelligible (or partially intelligible) speech signals by comparison with acoustically complex, unintelligible control conditions (Eisner et al., 2010; Evans et al., 2014; Narain et al., 2003; Scott et al., 2000) and by using parametric modulations of speech intelligibility, for example by varying the number of channels in noise-vocoded speech

(Davis & Johnsrude, 2003a, b; McGettigan et al., 2012b; Obleser et al., 2007; Scott et al., 2006). This work has identified that the process of extracting an intelligible message from an auditory speech signal engages an anterior-going pathway in the superior temporal lobes (Evans et al., 2014; Scott et al., 2000) as well as responses in the inferior frontal gyrus (IFG), anterior insula and premotor cortex (Adank, 2012a; Davis & Johnsrude, 2003a, b; Eisner et al., 2010; Hervais-Adelman et al., 2012; McGettigan et al., 2012b). Speech comprehension can also be manipulated experimentally through alternate methods, such as the comparison of words with pseudowords, and the use of semantic and syntactic violations, revealing similar loci (see Adank, 2012b). Although some authors argue that the perceptual processing of speech is bilateral in the temporal lobes (Hickok and Poeppel, 2007; Okada et al., 2010) our work has strongly suggested a left hemisphere dominance for intelligible speech perception (including perception of sentences, words, syllables, phonemes, syntactic and semantic information; see McGettigan and Scott, 2012), with a com-

* Corresponding author at: Department of Psychology, Royal Holloway, University of London, Egham Hill, Egham TW20 0EX, UK.
E-mail address: Carolyn.McGettigan@rhul.ac.uk (C. McGettigan).

<http://dx.doi.org/10.1016/j.neuropsychologia.2017.04.013>

Received 20 October 2016; Received in revised form 5 April 2017; Accepted 7 April 2017

Available online 08 April 2017

0028-3932/ © 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

plementary right-hemisphere dominance for the processing of melodic aspects of spoken signals and the perception of vocal identities (Scott et al., 2000; Kyong et al., 2014; McGettigan and Scott, 2012).

Previous investigations of audiovisual speech intelligibility have shown that the presence of dynamic facial cues improves speech report accuracy under difficult listening conditions (McGettigan et al., 2012a, b; Ross et al., 2007; Sumbly and Pollack, 1954). Neuroimaging studies of audiovisual speech perception have implicated sites including the posterior superior temporal sulcus (STS), inferior parietal cortex, motor cortex and subcortical structures such as the caudate nucleus (Bernstein et al., 2008; Calvert et al., 1997, 2001; McGettigan et al., 2012a, b; Skipper et al., 2005; Stevenson and James, 2009). However, beyond the basic cues to speech intelligibility from the movements of the lips and facial muscles, a talking face brings other information to a communicative interaction, including cues to mood and intentionality - salient amongst these cues is the gaze of the talker. Senju and Johnson (2009) consider the behavioural and neural effects of experiencing eye contact with another person. The authors identify a set of key brain regions that are regularly implicated in studies of gaze perception from faces, including the fusiform gyrus, anterior and posterior portions of the STS, medial prefrontal and orbitofrontal cortices, and the amygdala. They describe how perceived eye contact from another can increase autonomic arousal and modulate activation within the “social brain” (medial prefrontal cortex (mPFC), temporal poles and the temporoparietal junction (TPJ)), thus signaling communicative intent to this system. However, they also note inconsistencies in the neuroimaging literature on gaze perception, where some brain regions are only implicated across some studies, while other regions show contradictory responses from one study to the next (e.g. mPFC showing a preferential response to direct eye contact in one study, but to averted gaze in another). To make sense of these inconsistencies, Senju and Johnson propose their “fast-track modulator” model of eye gaze, in which they suggest that the fundamental mechanism for eye gaze detection is subcortical in its origin, and that the involvement of higher-order cortical centres is strongly dependent on task demands.

The STS has been a key region implicated in both the perceptual processing of both speech (Binder et al., 2000; Davis & Johnsrude, 2003a, b; Eisner et al., 2010; Evans et al., 2014; Liebenthal et al., 2005; Scott et al., 2000) and eye gaze direction (Calder et al., 2006, 2002; Carlin et al., 2011; Hoffman and Haxby, 2000; Hooker et al., 2003; Pelphrey et al., 2003, 2004; Puce et al., 1998). Senju and Johnson (2009) describe a difficulty in resolving the relative roles of posterior and anterior sites on the sulcus in terms of the response to eye contact and the mechanisms for discriminating gaze direction, where they suggest that directed attention to the eyes may be required to activate the anterior STS while posterior sites may require dynamic visual cues and/or conscious recognition of communicative intent from the viewed person. Calder and colleagues carried out a series of studies in which they attempted to better resolve mechanisms for gaze processing along the STS (Calder et al., 2006, 2002; Carlin and Calder, 2013; Carlin et al., 2011). In line with evidence from single-cell recordings in monkeys (e.g. Perrett et al., 1992; Perrett et al., 1985), Carlin et al. (2011) characterized an anterior-going processing hierarchy in the STS, where posterior sites are sensitive to both gaze and head direction while the anterior STS shows head-direction-invariant responses to gaze. In this way, it is suggested that the anterior STS is more responsive to the social significance of gaze than to the specific configurations of the visual cues that signal it. Here, there are strong parallels with the speech perception literature, in which there is a long-standing debate over whether the crucial mapping of sound to linguistic representations takes place primarily in posterior or anterior STS (Evans et al., 2014; Okada et al., 2010; Scott et al., 2000). There is an argument for a speech processing hierarchy in humans that is homologous to the ventral “what” pathway for auditory object recognition in the temporal lobe of non-human primates (Rauschecker and Scott, 2009; Scott and Johnsrude, 2003). The anterior STS forms the highest point in this

hierarchy, being responsive to intelligible speech signals regardless of their specific acoustic properties (e.g. whether they are undistorted or spectrally degraded; Scott et al., 2000). However, the main distinction between the findings with gaze and speech perception is one of hemispheric lateralization, where responses to speech tend to be left dominant in the superior temporal lobes (McGettigan and Scott, 2012), while sensitivity to gaze direction is more consistently *right-lateralized* (Carlin and Calder, 2013).

It is tautological to think of “social speech processing”, given that the vast majority of spoken language interactions take place in social settings. Nonetheless, the neurobiological literature has relatively little to say about social context for spoken communication in terms of how the processing of auditory information might interact with other social cues in speech perception (Scott et al., 2009). Above, we note the potential commonalities of anterior-going temporal lobe hierarchies for the processing of auditory speech and eye gaze direction. The STS has been repeatedly implicated in the processing of socially-relevant signals, including emotional prosody, facial expressions, vocal identity, gesture and biological motion (Allison et al., 2000; Belin et al., 2000; Grandjean et al., 2005; Grezes et al., 2003). In an attempt to unify this response profile in terms of its underlying computations, Redcay (2008) has proposed that the primary function of the STS is to parse and interpret the communicative significance of incoming streams of audio, visual and audiovisual information unfolding over time. Recent work using vocal signals supports this suggestion of the STS as a locus for social perception, finding that communicative speech and emotional vocalizations generated greater responses in the STS than non-communicative sneezes and coughs (Shultz et al., 2012). Similarly, the right posterior STS has been found to be specifically involved in the planning and perception of communicative (vs. non-communicative) actions in a two-player computer game (Noordzij et al., 2010; Stolk et al., 2013).

Given the apparent parallels between gaze and speech perception pathways, yet a strong difference in the reported lateralization of these processes, an unanswered question is whether and how the lateralization of one or both processes might be affected by the task context. Behaviourally, there is evidence that heard speech can influence the perception of gaze in a simultaneously presented face – participants were more likely to label gaze direction in a static face as direct when an accompanying voice called the participant's name (vs. a control name; Stoyanova et al., 2010). Kampe and colleagues (Kampe et al., 2003) presented participants with visual and auditory stimuli in separate trials of an event-related fMRI experiment. In the visual condition, they manipulated the gaze of static faces to be direct or averted with respect to the participant, while in the auditory condition a heard voice either demanded the attention of the participant by name (e.g. “Hey John!”) or addressed another person. Within each modality, the authors found modulation of key sites in the social brain by conditions with greater communicative intent (i.e. direct gaze and use of the participant's name) - the paracingulate cortex and the left temporal pole were implicated for both modalities. This is an important indication that there is sensitivity to gaze, and its social significance, in the left as well as the right temporal cortex. Using a region of interest analysis, Carlin et al. (2011) also reported head-view-invariant responses to gaze direction in the left anterior STS.

In Senju and Johnson's (2009) proposed “fast-track modulator” model of the eye contact effect, a subcortical stream forms the first path for detection of eye contact and projects to several sites in the social brain. These, in interaction with dorsolateral prefrontal responses to task demands and social context, then influence the further processing of gaze cues in the cortex. Although this model is focused on the specific percept yielded when another's eyes make direct contact with the gaze of the perceiver, its broad implication is that there is a dynamic network, or set of networks, underpinning the extraction of gaze cues in terms of their social and communicative significance. The aim of the current study was to explore the neural responses to talker gaze direction in the context of a speech intelligibility task. We had a particular

Download English Version:

<https://daneshyari.com/en/article/5045068>

Download Persian Version:

<https://daneshyari.com/article/5045068>

[Daneshyari.com](https://daneshyari.com)