NeuroImage 155 (2017) 97–112

Contents lists available at ScienceDirect

NeuroImage

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

# Developmental phonagnosia: Linking neural mechanisms with the behavioural phenotype

Claudia Roswandowitz[a,b,*], Stefanie Schelinski[a], Katharina von Kriegstein[a,c]

[a] Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany
[b] International Max Planck Research School on Neuroscience of Communication, Stephanstraße 1a, 04103 Leipzig, Germany
[c] Humboldt University zu Berlin, Rudower Chaussee 18, 12489 Berlin, Germany

ARTICLE INFO

*Keywords:*
Associative
Apperceptive
Identity
Functional MRI
Voice recognition

ABSTRACT

Human voice recognition is critical for many aspects of social communication. Recently, a rare disorder, developmental phonagnosia, which describes the inability to recognise a speaker's voice, has been discovered. The underlying neural mechanisms are unknown. Here, we used two functional magnetic resonance imaging experiments to investigate brain function in two behaviourally well characterised phonagnosia cases, both 32 years old: AS has apperceptive and SP associative phonagnosia. We found distinct malfunctioned brain mechanisms in AS and SP matching their behavioural profiles. In apperceptive phonagnosia, right-hemispheric auditory voice-sensitive regions (i.e., Heschl's gyrus, planum temporale, superior temporal gyrus) showed lower responses than in matched controls ($n_{AS}$=16) for vocal versus non-vocal sounds and for speaker versus speech recognition. In associative phonagnosia, the connectivity between voice-sensitive (i.e. right posterior middle/inferior temporal gyrus) and supramodal (i.e. amygdala) regions was reduced in comparison to matched controls ($n_{SP}$=16) during speaker versus speech recognition. Additionally, both cases recruited distinct potential compensatory mechanisms. Our results support a central assumption of current two-system models of voice-identity processing: They provide the first evidence that dysfunction of voice-sensitive regions and impaired connectivity between voice-sensitive and supramodal person recognition regions can selectively contribute to deficits in person recognition by voice.

## Introduction

The ability to recognise the identity of other people is critical for successful social interactions (von Kriegstein et al., 2008; Yardley et al., 2008, McGettigan, 2015). Impairments in person-identity processing lead to psychosocial disabilities such as difficulties in communication, avoidance of social situations, and feelings of embarrassment and failure (Yardley et al., 2008; Fine, 2012). A relatively large proportion of the population has selective developmental difficulties with recognising others by face (developmental prosopagnosia, McConachie, 1976; Behrmann and Avidan, 2005; Kennerknecht et al., 2006; Grueter et al., 2007). More recently, it has been discovered that also voice-identity processing can be selectively impaired without apparent brain lesion (developmental phonagnosia; Garrido et al., 2009a; Roswandowitz et al., 2014).

There is an ongoing debate about the neural deficits associated with developmental prosopagnosia (Avidan and Behrmann, 2009; Garrido et al., 2009b; Furl et al., 2010). The underlying neural mechanisms for developmental phonagnosia are even less clear (Xu et al., 2015). The first aim of the present study was to investigate the neural mechanisms for developmental phonagnosia. The second aim was to test a central prediction of current neuroscientific models of person recognition.

The standard neuroscientific models of person recognition contain a so-called core and an extended system, which are connected with each other (Haxby et al., 2000; Belin et al., 2004; Gobbini and Haxby, 2007; Avidan and Behrmann, 2014; Blank et al., 2014; Gainotti, 2015; Rice et al., 2015). The core system processes modality-specific perceptual and identity information of the face or voice (Kanwisher et al., 1997, Belin et al., 2000, von Kriegstein et al., 2003, Garrido et al., 2009b, Furl et al., 2010, Bernstein and Yovel, 2015). The extended system encodes multi-modal semantic information about a person, such as the occupation or the name (Damasio et al., 2004; Patterson et al., 2007; Visser et al., 2009; Simmons et al., 2010; Gainotti, 2013; Abel et al., 2015). The two-system architecture with a core and an extended system was originally proposed for face-identity processing (Haxby et al., 2000), but it is also implicit in models for voice-identity processing (Belin et al., 2004; Blank et al., 2014; Perrodin et al., 2015). A central prediction of the two-system models of person recognition is that modality-specific person-identity processing deficits can either be caused by a dysfunction in the core system or by a disconnection

* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany.
*E-mail address:* roswandowitz@cbs.mpg.de (C. Roswandowitz).

http://dx.doi.org/10.1016/j.neuroimage.2017.02.064
Received 15 June 2016; Accepted 21 February 2017
Available online 22 February 2017

between the core and the extended system. Our aim was to test this central hypothesis in developmental phonagnosia.

Worldwide there are currently four reported cases of developmental phonagnosia (Garrido et al., 2009b; Roswandowitz et al., 2014; Xu et al., 2015). Here, we applied functional magnetic resonance imaging (fMRI) in two behaviourally well characterised cases: AS and SP (Roswandowitz et al., 2014). AS, a 32-year-old female, has apperceptive phonagnosia: Her deficit in voice-identity processing is characterised by difficulties with the analysis and integration of acoustic voice features. The association of semantic information to the voice is intact. SP, a 32-year-old male, has associative phonagnosia: His deficit in voice-identity processing is characterised by difficulties in associating semantic information to the voice, while the perception of the voice is intact (Roswandowitz et al., 2014). Cases with developmental person-recognition impairments provide a unique opportunity to test specific predictions made by neuroscientific models of person recognition, because of the selectiveness of the behavioural impairment and the absence of brain damage.

Most of what we currently know about how the brain processes voices comes from neuroimaging studies on healthy, i.e. neurotypically developed populations. Candidate areas for the core-voice system are located in the auditory cortex, i.e., the Heschl's gyrus (Formisano et al., 2008, Bonte et al., 2014), the planum temporale (von Kriegstein and Giraud, 2006; Warren et al., 2006), and several regions in the superior and middle temporal gyrus (STG/MTG) and superior temporal sulcus (STS) (e.g., Belin et al., 2000; von Kriegstein and Giraud, 2004; Warren et al., 2006; Pernet et al., 2015). Voice-identity processing also leads to responses in candidate areas for an extended system: supramodal brain regions such as the precuneus/posterior cingulate, the temporal pole, or the amygdala, inferior frontal gyrus, and regions of other sensory modalities, such as the fusiform face area (Shah et al., 2001; von Kriegstein et al., 2005; von Kriegstein and Giraud, 2006; Andics et al., 2010; Latinus et al., 2011; Schall et al., 2013, for review Blank et al., 2014). These regions are functionally or/and structurally connected to regions of the core-voice system (von Kriegstein et al., 2005; von Kriegstein and Giraud, 2006; Blank et al., 2011; Blank et al., 2014).

We used two standard fMRI paradigms to elicit responses in the core-voice and the extended system: a vocal-sound experiment (Belin et al., 2000) and a speaker-identity experiment (von Kriegstein et al., 2003). In the vocal-sound experiment, participants passively listened to vocal and non-vocal sounds (Belin et al., 2000). In the speaker-identity experiment, participants either performed a speaker or a speech task on sentences spoken by different speakers (adapted from von Kriegstein et al., 2003; Blank et al., 2011; Schelinski et al., 2016). We interpreted decreased blood oxygen level dependent (BOLD) response in the phonagnosia cases in comparison to controls as reflecting a malfunctional region. Conversely, we interpreted higher BOLD response in the phonagnosia cases compared to controls as an attempt to compensate for the behavioural deficit. This is commonly done in fMRI studies on developmental deficits (Dricot et al., 2008; Kaiser et al., 2010; Hoeft et al., 2011; Avidan and Behrmann, 2014). Because of the distinct behavioural profiles of AS and SP, we predicted two distinct profiles in the fMRI results. For apperceptive phonagnosia (AS), we expected dysfunction of the core-voice system, while the extended system responses remain intact or are even elevated as an attempt to compensate. In contrast, associative phonagnosia (SP) should be associated with dysfunctional connectivity between the core-voice and the extended system and intact or even increased BOLD response in the core-voice system as a compensatory attempt.

## Materials and methods

### *Participants*

#### *Cases AS and SP*

AS, a 32-year-old female graduate student, and SP, a 32-year-old male PhD student, both reported life-long voice-recognition deficits. We identified both participants via a four-stage screening approach and characterised their voice-identity processing deficits with a comprehensive behavioural test battery in a previous study (Roswandowitz et al., 2014). The behavioural test battery included tests on unfamiliar voice discrimination, unfamiliar voice learning and recognition (voice-face, voice-colour, and voice-name association learning), familiar voice recognition, and tests on vocal-pitch and vocal-timbre discrimination. Further it comprised several control tests on other auditory skills (e.g., speech, vocal-emotion, and music processing) and visual person-identity processing (e.g. unfamiliar face-identity recognition). For detailed test descriptions please refer to Roswandowitz et al. (2014).

Both cases were severely impaired in learning and recognising unfamiliar voices when the learning involved associating a name, face, or colour to the unfamiliar voices. What distinguished both cases were the performances on the unfamiliar voice discrimination test and the familiar voice recognition test. AS was impaired in discriminating unfamiliar voices, i.e. judging whether two voice samples were from the same or different person. Conversely, the familiar voice recognition test showed that her ability to associate semantic information to voices that she had correctly classified as familiar was intact. We had therefore classified her as a case of apperceptive phonagnosia (Roswandowitz et al., 2014) in analogy to similar classifications of other auditory and visual agnosias (Lissauer, 1890; Vignolo, 1969; Buchtel and Stewart, 1989; De Renzi et al., 1991; Griffiths et al., 1999). In contrast, SP's performance in discriminating one unfamiliar voice from the other (unfamiliar voice discrimination test) was normal. Conversely, his ability to associate semantic information to voices that he had correctly recognised as familiar (familiar voice recognition test) was impaired. We had therefore classified him as a case of associative phonagnosia (Roswandowitz et al., 2014). Both cases also had an elevated threshold for discriminating vocal pitch (for a discussion how this might relate to developmental phonagnosia, see Roswandowitz et al., 2014). The impairments of AS and SP were restricted to voice-identity and vocal-pitch processing, their performances in all other control tasks were comparable to controls.

Both cases did not have a history of psychiatric or neurological disease, or developmental deficits, with the exception of developmental phonagnosia.

#### *Control participants*

We tested in total 32 control participants, 16 matched controls for AS and SP respectively (AS's controls: all female, mean age=31.63 years, SD=1.22, range=30–34 years; SP's controls: all male, mean age=31.06 years, SD=1.2, range=29–3 years). Control participants were matched to AS and SP for the following criteria: age, gender, education, and handedness (Edinburgh questionnaire, Oldfield, 1971).

To assess voice-identity processing abilities among control participants, they took part in a behavioural test battery. All controls were assessed on unfamiliar voice discrimination and speaker recognition, both tests were part of the fMRI speaker-identity experiment. Further we tested controls (AS's controls $n$=15, SP's controls $n$=12) on their ability to learn and recognise unfamiliar voices with a voice-face and voice-colour test (for test descriptions see Roswandowitz et al., 2014). If a participant performed significantly below the control's mean in more than one test, we considered this participant as potentially phonagnosic. None of the control participants had test performances, which were indicative of phonagnosia.

None of the control participants reported a history of neurological or psychiatric disease. Further, both phonagnosia cases and control participants reported to have normal or corrected to normal vision and had normal hearing levels. We formally assessed hearing levels via pure-tone audiometry (250–8000 Hz) using a screening audiometry (MADSEN Micromate 304, GN Otometrics, Copenhagen, Denmark).