NeuroImage 134 (2016) 12–21

Contents lists available at ScienceDirect

NeuroImage

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

Neurobiology of knowledge and misperception of lyrics

Claudia Beck Lidén [a], Oliver Krüger [a], Lena Schwarz [a,b], Michael Erb [a], Bernd Kardatzki [a], Klaus Scheffler [a,c], Thomas Ethofer [a,b,c,*]

[a] Department of Biomedical Magnetic Resonance, University of Tübingen, Otfried-Müller-Str. 51, 72076 Tübingen, Germany
[b] University Clinic for Psychiatry and Psychotherapy, University of Tübingen, Calwer Str. 14, 72076 Tübingen, Germany
[c] Max-Planck-Institute for Biological Cybernetics, Speemannstraße 38-40, 72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 16 November 2015
Revised 30 March 2016
Accepted 31 March 2016
Available online 13 April 2016

Keywords:
Connectivity
DWI
fMRI
Lyrics
Misperception
Music

ABSTRACT

We conducted two functional magnetic resonance imaging (fMRI) experiments to investigate the neural underpinnings of knowledge and misperception of lyrics. In fMRI experiment 1, a linear relationship between familiarity with lyrics and activation was found in left-hemispheric speech-related as well as bilateral striatal areas which is in line with previous research on generation of lyrics. In fMRI experiment 2, we employed so called Mondegreens and Soramimi to induce misperceptions of lyrics revealing a bilateral network including middle temporal and inferior frontal areas as well as anterior cingulate cortex (ACC) and mediodorsal thalamus. ACC activation also correlated with the extent to which misperceptions were judged as amusing corroborating previous neuroimaging results on the role of this area in mediating the pleasant experience of chills during music perception. Finally, we examined the areas engaged during misperception of lyrics using diffusion-weighted imaging (DWI) to determine their structural connectivity. These combined fMRI/DWI results could serve as a neurobiological model for future studies on other types of misunderstanding which are events with potentially strong impact on our social life.

Introduction

"Gib mir die Hand, ich geb Dir die Meine" (give me your hand, I give you mine) was the refrain of a church hymn played during a wedding ceremony, but the bridegroom repeatedly understood "Gib mir die Hand, ich geb Dir die Beine" (give me your hand, I give you the legs) and was wondering about the meaning of these words in the context of his marriage. This is a typical example on how perception can fail due to expectations (in this case generated by the fact that hand and leg belong both to the semantic word field of body parts). Generation of predictions is a fundamental mechanism of sensory systems to maximize the conveyed information while minimizing the required resources which has been defined within the concept of predictive coding (Huang and Rao, 2011). Removal of predictable components by learning statistical regularities is a highly efficient way of coding sensory input, but can also result in perception which does not accurately reflect the physical stimulus. While in this case the misperception was clearly harmless and its enlightenment a reason for joy, misunderstandings caused by such extrapolations of predictive coding mechanisms can have a severe impact on one's social life. It is difficult, however, to examine such phenomena within the framework of a neuroimaging experiment as this requires both to control induction of wrong expectations and the onset of their effects on speech perception. To overcome this methodological obstacle we recently established a method to induce misperceptions on the basis of misheard lyrics (Beck et al., 2014). Lyrics represent speech stimuli which are typically perceived under adverse conditions due to often atypical pronunciation and background noise (instrumental music) and such suboptimal conditions decrease intelligibility (Mattys et al., 2012). Misperceptions within the language of the original lyrics are called "Mondegreen" in reference to their first description by Sylvia Wright for the Scottish folk song "The Bonny Earl O'Morray" (Wright, 1954) in which the author understood in its last line "They have slain Earl O'Morray/and Lady Mondegreen" instead of "They have slain the Earl O'Morray/and laid him on the green". Here, the mishearing is based on wrong segmentation of continuous speech into its word components, for a review of this phenomenon see Cutler and Butterfield, 1992). Non-native stimuli are particularly prone to perceptual illusions including 'nativization processes' (Calabrese, 2012). If mishearing of lyrics results in homophonic/near homophonic translations across languages (typically into the native language of the listener), they are called "Soramimi" (which means mishearing in Japanese where the observation of this phenomenon has its longest tradition). It has been shown that concurrent presentation of written text can serve as a top–down modulation enhancing the clarity of speech stimuli (Sohoglu et al., 2014). In the present study, we make use of this phenomenon to induce Mondegreens/Soramimi which enables us to experimentally control both the expectation as well as the exact onset of the resulting misperception.

* Corresponding author at: Department of Biomedical Magnetic Resonance, University of Tübingen, Otfried-Müller-Str. 51, 72076 Tübingen, Germany.

http://dx.doi.org/10.1016/j.neuroimage.2016.03.080
1053-8119/

We conducted two functional magnetic resonance imaging (fMRI) experiments. In fMRI experiment 1, we aimed to clarify which neural structures underlie knowledge of lyrics. We measured knowledge of lyrics by asking the participants to rate their familiarity with the lyrics of song excerpts on a 4-point scale presented during scanning. A previous neuroimaging study (Janata, 2009) comparing familiar with unfamiliar song excerpts revealed activations in the left inferior frontal cortex and posterior superior temporal gyrus which have been associated with speech-related processes such as word retrieval and speech comprehension (for a review, see Price, 2012). On the other hand, adaptation effects due to repeated exposure of lyrics resulted in bilateral repetition suppression effects in the superior temporal cortex (Sammler et al., 2010).

Before fMRI experiment 2, we informed the study participants about ambiguous parts in these song excerpts with the intention to systematically induce misperceptions (i.e. perception of alternative lyrics not intended by singer, but made possible by the ambiguity of the acoustic information). Regarding the neural correlates of such altered speech perception, we based our hypothesis on the available neuroimaging literature on bistable auditory stimuli. This approach induces two different percepts in spite of physical stimulus constancy and has been repeatedly employed using visual stimuli (e.g. Rubin's face-vase illusion, Andrews et al., 2002; Hasson et al., 2001). In the auditory domain, first results based on simple tone stimuli to induce the octave illusion (Deutsch, 1974) point to a bilateral involvement of auditory and frontal cortices associated with the two concurring percepts (Brancucci et al., 2011). So far, little is known regarding the representation of such competing perceptions in more complex stimuli, such as lyrics. It should be noted, however, that a network specifically engaged to induce misperceptions in speech is not biologically plausible as it would offer no advantage from an evolutionary perceptive. Thus, the aim of this study is not to propose a 'misperception network', but rather clarify which parts of the regular speech perception network fail during misperceptions induced by Mondegreens/Soramimi. A dual-stream speech perception model has been proposed with a predominantly left-lateralized dorsal stream for auditory motor-integration and a bilaterally organized ventral stream for retrieval of meaning from sounds (Hickok, 2012; Hickok and Poeppel, 2000, 2007). As misheard lyrics are associated with a change in meaning, the ventral stream regions in bilateral superior/middle temporal gyrus are therefore potential candidates mediating induced misperceptions of speech stimuli. Regarding processing of music, comparison of sung versus spoken sentences revealed an overlapping, but differential representation within the temporal lobe (Schön et al., 2010) as well as the inferior frontal cortex (Merrill et al., 2012) and these areas also showed differential activation in the "song versus speech auditory illusion" (Tierney et al., 2013). The critical results of these studies are, however, dependent on activation differences to two types of stimuli (sung versus spoken words) and thus on a bottom-up effect. In the current study, we aimed to test whether these areas are also subject to top–down modulation (i.e. a differential activation to the same physical stimuli driven before versus after experimentally induced misperceptions). Moreover, as such misperceptions reflect changes of the perceived semantic meaning of lyrics, we hypothesized that altered activation partially overlaps with brain areas associated with knowledge of lyrics as determined in fMRI experiment 1.

Materials and methods

Ethics statement

The study was approved by the Ethical Committee of the University of Tübingen (votum: 215/2012BO2) and written informed consent was obtained from all participants. All study procedures were in line with the latest version of the Declaration of Helsinki.

Participants

20 healthy German native speakers (9 women, 11 men, mean age: 28.9 ± 5.4 years, education: 16.8 ± 2.6 years) participated in this study. All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Mean verbal intelligence of the participants, as obtained by a German vocabulary test (Mehrfach-Wortschatz-Intelligenztest B), was 123.8 ± 13.6. Knowledge of English language was assessed using a language test based on a short form of the Test of English as a Foreign Language (Mini-TOEFL, http://www.ets.org/toefl) to determine the participants' grammar and vocabulary comprehension. On average, the participants correctly answered 14.7 ± 5.4 out of 25 questions. Verbal intelligence and English language comprehension were significantly correlated ($r = 0.53$, two-tailed $p < 0.05$) reflecting the fact that they are similarly driven by the educational level of the participants. All values are given in mean ± standard deviation.

Stimuli

The stimulus set comprised short audio clips (mean duration: 17.7 ± 3.0 s) taken from 18 English and 18 German songs. These 36 songs were selected from a set of 41 English and 20 German songs which were evaluated in a behavioral experiment (Beck et al., 2014) to ensure that both types of stimuli were matched for degree of familiarity with the lyrics, strength of the induced misperception, and wittiness of the alternative lyrics as evaluated on a four-point scale as well as prior knowledge of the alternative lyrics as determined on a two-alternative forced choice decision (see Table 1). No significant differences were found for any of these stimulus characteristics (all $p > 0.10$). All stimuli were normalized to same peak intensity.

Experimental design

All stimuli were presented during four consecutive experimental runs in a fully randomized order. Experimental run 1 and 4 were carried out during fMRI while run 2 and 3 were conducted outside the scanner as behavioral experiments. Directly before each of the four runs, the participants were familiarized with the task by a short training run consisting of 2 stimuli which were not used during the experimental runs. The same two training stimuli were used in all four training runs. In all four experimental runs, the response scales were shown for five seconds and followed by a fixation cross presented for two seconds before onset of the next stimulus. During experimental run 1 (fMRI experiment 1) the participants judged their degree of familiarity with the lyrics on a four-point scale (+++: 'I know most/all of the lyrics', ++: 'I know some of the lyrics in addition to the refrain', +: 'I only know the refrain', 0: 'I do not even know the full refrain'). Before starting experimental run 2 (behavioral experiment 1), the participants were informed that the lyrics can be misheard and that such alternative lyrics will be presented visually on the screen while listening again to the same stimuli in the next two experimental runs. The task of the participants in experimental run 2 was to determine whether these alternative lyrics were already known to them (e.g., because they heard them in radio transmissions or even spontaneously misheard the lyrics in the presented manner) on a two-alternative scale. In experimental

Table 1
Characteristics of employed stimulus material.

	English songs	German songs
Familiarity with lyrics	1.44 ± 0.39	1.16 ± 0.81
Strength of induced misperceptions	1.94 ± 0.48	1.90 ± 0.53
Wittiness of alternative lyrics	1.27 ± 0.42	1.23 ± 0.37
Prior knowledge of alternative lyrics	16% ± 16%	9% ± 10%

All values in mean ± standard deviation.