Brain and Cognition 91 (2014) 35–44

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Silent music reading: Auditory imagery and visuotonal modality transfer in singers and non-singers

Christian Hoppe [a,1,*], Christoph Splittstößer [a,1], Klaus Fliessbach [c,d], Peter Trautner [b], Christian E. Elger [a,b], Bernd Weber [a,b,e]

[a] Department of Epileptology, University of Bonn Medical Centre, Germany
[b] Life & Brain Center, University of Bonn Medical Centre, Germany
[c] Department of Psychiatry, University of Bonn Medical Centre, Germany
[d] German Center for Neurodegenerative Diseases, Germany
[e] Center for Economics and Neuroscience, University of Bonn, Germany

A R T I C L E I N F O

Article history:
Accepted 10 August 2014
Available online 15 September 2014

Keywords:
Working memory
Mental imagery
Modality transfer
Musical expertise
Music reading
Functional magnetic resonance imaging

A B S T R A C T

In daily life, responses are often facilitated by anticipatory imagery of expected targets which are announced by associated stimuli from different sensory modalities. Silent music reading represents an intriguing case of visuotonal modality transfer in working memory as it induces highly defined auditory imagery on the basis of presented visuospatial information (i.e. musical notes). Using functional MRI and a delayed sequence matching-to-sample paradigm, we compared brain activations during retention intervals (10 s) of visual (VV) or tonal (TT) unimodal maintenance versus visuospatial-to-tonal modality transfer (VT) tasks. Visual or tonal sequences were comprised of six elements, white squares or tones, which were low, middle, or high regarding vertical screen position or pitch, respectively (presentation duration: 1.5 s). For the cross-modal condition (VT, session 3), the visuospatial elements from condition VV (session 1) were re-defined as low, middle or high "notes" indicating low, middle or high tones from condition TT (session 2), respectively, and subjects had to match tonal sequences (probe) to previously presented note sequences. Tasks alternately had low or high cognitive load. To evaluate possible effects of music reading expertise, 15 singers and 15 non-musicians were included. Scanner task performance was excellent in both groups. Despite identity of applied visuospatial stimuli, visuotonal modality transfer versus visual maintenance (VT > VV) induced "inhibition" of visual brain areas and activation of primary and higher auditory brain areas which exceeded auditory activation elicited by tonal stimulation (VT > TT). This transfer-related visual-to-auditory activation shift occurred in both groups but was more pronounced in experts. Frontoparietal areas were activated by higher cognitive load but not by modality transfer. The auditory brain showed a potential to anticipate expected auditory target stimuli on the basis of non-auditory information and sensory brain activation rather mirrored expectation than stimulation. Silent music reading probably relies on these basic neurocognitive mechanisms.

> *"I believe that the experience of having an image is just the inner aspect of a readiness to perceive the imagined object, and that differences in the nature and quality of people's images reflect differences in the kind of information they are prepared to pick up."*
> Ulric Neisser (1928–2012), from: Cognition and Reality (1976, p. 130f.)

1. Introduction

While inner mental life hitherto appeared essentially private and inaccessible to empirical research, subjective reports of inner experience can nowadays be objectively cross-checked to some degree by functional neuroimaging on the basis of correlated brain activation patterns (Kosslyn, Ganis, & Thompson, 2001; McNorgan, 2012; Zatorre, 2003; Zatorre & Halpern, 2005). Notational audiation, i.e. the inner experience of music in the absence of acoustic input as induced by silent music reading, provides one of the most intriguing examples of mental imagery (Godøy & Jørgensen, 2001) and appears well suited for neurocognitive research as the score precisely defines the music to be imagined (e.g. rhythm, pitch,

* Corresponding author. Postal address: Department of Epileptology, University of Bonn Medical Center, Sigmund-Freud-Strasse 25, FRG-53105 Bonn, Germany. Fax: +49 228 287 90 16172.

E-mail address: christian.hoppe@ukb.uni-bonn.de (C. Hoppe).

[1] *Note*: First authorship is shared by C. Hoppe and C. Splittstößer.

http://dx.doi.org/10.1016/j.bandc.2014.08.002
0278-2626/

harmonies, sound, and emotional quality). Meanwhile, a broad stream of neurocognitive findings have corroborated the notion of subjective auditory experience during music imagery by showing concomitant activation of auditory brain areas in the absence of acoustic input (Brodsky, Kessler, Rubinstein, Ginsborg, & Henik, 2008; Gordon, 1975; Halpern, 2001; Hubbard, 2010; Zatorre, 2003; Zatorre & Halpern, 2005).

From a working memory perspective, silent music reading translates into a case of visuospatial-to-tonal modality transfer in working memory (Calvert, 2001; Evans & Treisman, 2010). Thereby, modality transfer is supposed to be a latent cognitive mechanism which facilitates later responses to target stimuli which were previously announced by associated stimuli from a different modality (Calvert, 2001). Visuo-tonal modality transfer provides an opportunity to test contradictory implications of the two most influential working memory models, the *multi-component model* (Baddeley & Hitch, 1974; Repovš & Baddeley, 2006) and the *embedded-processes model* (Cowan, 1999; Cowan, Sossin, Lacaille, Castellucci, & Belleville, 2008). Of note, tonal material can largely be treated like other phonological material in this context (Akiva-Kabiri, Vecchi, Granot, Basso, & Schön, 2009; Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011; Williamson, Baddeley, & Hitch, 2010). While both accounts appear consistent with a visual-to-auditory brain activation shift during silent music reading, they make different predictions regarding motor processes. As the multi-component model stresses the modality specificity of the storage units it essentially requires (subvocal) articulatory processes to transfer visually presented material into a phonological code before it may enter the phonological loop (Muller & Knight, 2006; Repovš & Baddeley, 2006). In contrast, the embedded-processes model predicts a shift of the focus of attention from visual to auditory aspects of given information, but no motor processes. Thus, the activation or non-activation of speech-related motor brain areas (e.g. supplementary-motor area, SMA) during visuotonal modality transfer becomes a critical marker for the validity of both accounts.

In this functional magnetic resonance imaging (fMRI) study, we assessed neural mechanisms behind auditory imagery as induced by silent music reading, i.e. visuotonal modality transfer. We predicted that visuotonal modality transfer is associated with a shift of activation from visual to auditory brain areas during the retention interval of a delayed matching-to-sample task. More precisely, we tested the following hypotheses: (1) Visual stimuli rapidly gain a potential to activate auditory brain areas if they are associated with tonal information by re-defining them as notes indicating pitch. (2) Activation of auditory brain areas by visually presented notes in a cross-modal task might be even stronger than activation of auditory cortices elicited by tones in a unimodal maintenance task due to the additional requirement of self-generating a tonal representation during modality transfer (Suchan, Linnewerth, Koster, Daum, & Schmid, 2006). (3) In the absence of task-relevant acoustic input, visual stimulation might not only activate secondary and higher but even the primary auditory cortex (McNorgan, 2012). (4) The proposed visual-to-auditory activation *shift* also implies "deactivation" or inhibition of visual brain areas during increased auditory activation (Azulay, Striem, & Amedi, 2009; Hairston et al., 2008; Laurienti et al., 2002), consistent with the notion of highly automated visual processing but almost complete absorption in *musical* imagery during score reading in experts. (5) Activation of motor brain areas during visual-phonological modality transfer might indicate executive processes as predicted by the multi-component model of working memory. (6) We included singers with high levels of sight-singing skills and unskilled lay persons to test the hypothesis that expertise in terms of extended practice and higher performance in visuotonal modality transfer affects the proposed visual-to-auditory brain activation shift during silent music reading. Overall, we hypothesized that imagery might indicate an anticipatory role of sensory brain areas in response facilitation, beyond mere reception and storage of given information.

2. Methods

This study was carried out in accordance with *The Code of Ethics of the World Medical Association* (2008) for experiments involving humans and approved by the Ethical Review Board of the Medical Faculty at the University of Bonn (No. 245/09).

2.1. Subjects

We included 15 singers as experts and 15 non-musicians as controls matched for age (singers: $M = 25.6$, $SD = 2.7$ yrs.; controls: $M = 25.9$, $SD = 2.1$ yrs.) and gender (7 males/8 females in both groups). Experts were either professional singers ($n = 1$), singing students currently studying at local conservatories ($n = 7$) or ambitioned amateurs who had previously taken vocal lessons and regularly practice singing in a choir that performs on a high level ($n = 7$). All subjects had to have normal hearing ability and normal or corrected-to-normal vision. Subjects were reimbursed for travel costs and 10€ per hour for participation.

2.2. Study design and procedure

We assessed the effect of two factors on behavioral performance (i.e. recognition accuracy) and regional brain activation patterns: silent music reading expertise (experts, non-experts) and task condition, i.e. the combination of stimulus and probe modality in a delayed sequence matching task (unimodal visual maintenance, unimodal tonal maintenance, and cross-modal visuotonal modality transfer).

After finishing the informed consent process, subjects directly underwent the scanning procedure (see below). To avoid any association or suggestion with regard to music or notes before performance of the silent music reading task condition, personal data including the history of music training were captured only after scanning. The assessment was completed by adjunctive tests and ratings on non-verbal intelligence and the musical skill level (see below). The total duration of the examination remained below two hours in all subjects.

2.3. Cognitive activation task

The applied paradigm is shown in Fig. 1. In three subsequent sessions, subjects had to complete delayed matching-to-sample tasks on visually or auditorily presented sample and probe sequences (software: Presentation©/NeuroBehavioural Systems Inc.). The duration of the retention interval was 10 s during which a fixation cross was displayed on the screen center.

In the first session, subjects were introduced to a delayed unimodal sequence matching-to-sample task with sequences comprised of simple visuospatial elements as stimuli and probes (VV). The sequences consisted of 6 small white squares (17 × 17 pixels) which were serially presented from left to right at a low, medium or high vertical position relative to a presented working square. The visuospatial material was presented via video goggles (Nordic NeuroLab, Bergen, Norway; monitor resolution: 1024 × 768 pixels). Presented squares remained on screen until they disappeared altogether after presentation of the last square. The time between the emergence of single squares was set to 250 ms (equating 120 beats/minute; sequence presentation duration: 1.5 s). To avoid any effects of priming towards music or notes