Neural correlates of binding lyrics and melodies for the encoding of new songs

Irene Alonso [a,b,c,e], Lila Davachi [e,f], Romain Valabrègue [c,g], Virginie Lambrecq [b,c,d,g], Sophie Dupont [b,c,g], Séverine Samson [a,b,*]

[a] Laboratoire PSITEC (EA 4072), Université de Lille, France
[b] Epilepsy Unit, Hôpital de la Pitié-Salpêtrière, Paris, France
[c] Inserm U 1127, CNRS UMR 7225, Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Institut du Cerveau et de la Moelle épinière — ICM, F-75013 Paris, France
[d] IHU-ICM, F-75013 Paris, France
[e] Department of Psychology, New York University, New York, USA
[f] Center for Neural Science, New York University, New York, USA
[g] Centre de Neuroimagerie de Recherche — CENIR, F-75013 Paris, France

ARTICLE INFO

ABSTRACT

Songs naturally bind lyrics and melody into a unified representation. Using a subsequent memory paradigm, we examined the neural processes associated with binding lyrics and melodies during song encoding. Participants were presented with songs in two conditions: a unified condition (melodies sung with lyrics), and a separate condition (melodies sung with the syllable "la"). In both cases, written lyrics were displayed and participants were instructed to memorize them by repeating them covertly or by generating mental images of the songs. We expected the unified condition to recruit the posterior superior temporal gyrus, known to be involved in perceptual integration of songs, as well as the left inferior frontal gyrus (IFG). Conversely, we hypothesized that the separate condition would engage a larger network including the hippocampus to bind lyrics and melodies of songs, and the basal ganglia and the cerebellum to ensure the correct sequence coupling of verbal and musical information in time. Binding lyrics and melodies in the unified condition revealed activation of the left IFG, bilateral middle temporal gyrus (MTG), and left motor cortex, suggesting a strong linguistic processing for this condition. Binding in the separate compared to the unified condition revealed greater activity in the right hippocampus as well as other areas including the left caudate, left cerebellum, and right IFG. This study provides novel evidence for the role of the right hippocampus in binding lyrics and melodies in songs. Results are discussed in light of studies of binding in the visual domain and highlight the role of regions involved in timing and synchronization such as the basal ganglia and the cerebellum.

© 2015 Published by Elsevier Inc.

Introduction

Although enjoyable, widespread, and seemingly effortless, learning and remembering songs entail high-level cognitive processes that require the binding of verbal and musical information into a unified representation. Neuroimaging studies of song perception have led to enlightening debates on the independence or interaction of music and language networks (Schön et al., 2005). Two studies have suggested the involvement of the posterior areas of the superior temporal gyrus (STG) and superior temporal sulcus (STS) in the integration of lyrics and melodies within songs (Schön et al., 2010; Sammler et al., 2010). Despite these few studies that investigated the perceptual integration of verbal and musical information within songs, the neural mechanisms supporting the binding of lyrics and melodies in memory have received relatively little attention. To further explore this issue, we examined the cerebral structures underlying the binding of lyrics and melodies during encoding of new songs.

One widely used approach to study successful memory formation has been the subsequent memory paradigm (Wagner, 1998; Kim, 2011; Paller and Wagner, 2002). This paradigm allows identifying brain activity during encoding that can predict later success in recognition by comparing subsequently recognized trials with subsequently forgotten trials, known as the subsequent memory effect (SME). Such a paradigm has been instrumental in revealing the distinct contributions of regions in the medial temporal lobe (MTL) to memory (Davachi et al., 2003; Staresina and Davachi, 2006; Staresina and Davachi, 2009; Staresina et al., 2011). One recent study used the subsequent memory paradigm to examine the binding of visual objects features (i.e. shape and colors) in memory (Staresina and Davachi, 2009). In one condition, objects images were presented with a color in a unified

* Corresponding author at: Laboratoire PSITEC (EA 4072), Université de Lille, Pont de Bois BP 60149 F-59 653 Villeneuve d'Ascq Cedex, France. Tel.: +33 320416443.
E-mail address: severine.samson@univ-lille3.fr (S. Samson).

manner. In two other conditions, gray-scale objects and color were presented separately. The object in gray-scale was presented with a color frame displayed either simultaneously with the object (spatial discontinuity condition) or with a short delay (spatial–temporal discontinuity condition). In these discontinuity conditions, participants were instructed to use imagery to bind these features into a unified representation. Interestingly, the magnitude of the SME in the hippocampus, unlike the neighboring perirhinal cortex, increased as a function of the spatiotemporal discontinuity of the presented object and color representations.

To further explore the role of the hippocampus in memory binding, we developed a new paradigm involving auditory information. In this case, participants were required to bind lyrics with their associated melodies during song encoding. Hence, the present study aims at exploring the underlying brain mechanisms that support the binding of lyrics and melodies within songs. We examined subsequent memory for songs under two encoding conditions inspired by Staresina and Davachi's study (2009): 1) following a unified presentation of song components (sung lyrics) and 2) following a separate presentation of song components (simultaneous presentation of sung melody on the syllable "la" and of written lyrics). In both conditions, participants were asked to covertly and repeatedly sing the resulting song until the next trial and to retain the song for a recognition test. Critically, these two conditions are inherently different in terms of the cognitive effort of the participant. The unified presentation is a perceptually richer condition that already provides an integrated signal while the separate condition requires the participant to actively create a mental representation of the song.

Taking in consideration memory models describing the MTL binding function (Davachi, 2006; Diana et al., 2007; Ranganath, 2010) as well as recent empirical findings in the visual domain (Staresina and Davachi, 2009), we suggest that the hippocampus may be required to integrate various elements of a song into a unified memory trace. Based on results reported by Staresina and Davachi (2009), we suggest that the involvement of the hippocampus during encoding of songs will enhance when melody and lyrics are presented separately rather than in a unified presentation (sung lyrics), the hippocampus being particularly important to actively integrate separate components of an event.

As previously discussed, binding is required to integrate information both across space and time (Eichenbaum, 2013). It is therefore worth noting that these dimensions might interact differently in the visual and the auditory domains. On one hand, vision is largely governed by space, requiring rapid binding of synchronous events (Engel et al., 1997). On the other hand, audition is constrained by time variations (Sanders et al., 2011; Kiebel et al., 2008; Demany et al., 2010) and consequently requires precise timing and binding of unfolding sequences. Thus, the evoked binding network cooperating with the hippocampus may vary according to spatial and temporal demands of the task. In particular, the cerebellum and the basal ganglia, which are highly interconnected (Bostan et al., 2013), might be critical for binding auditory information. A non-motor role of the basal ganglia on speech perception has been previously proposed (for a review see Kotz et al., 2009; Kotz and Schwartze, 2010), and cerebellum activations have been frequently found during song perception and production (for a review see Callan et al., 2007). From a more general perspective both the cerebellum and basal ganglia have been shown to be crucial for sequence learning (Graybiel et al., 1994; Kotz et al., 2009; Shin and Ivry, 2003), and for generating temporal (Cope et al., 2014; Penhune et al., 1998; Schwartze and Kotz, 2013; for a review see Schwartze et al., 2012) and harmonic expectations (Seger et al., 2013). It is possible that activation of the cerebellum and basal ganglia may allow temporal coupling of syllables with musical notes during covert rehearsal, enhancing learning of the song sequence through auditory–motor binding. This idea is supported by studies reporting striatal contributions to auditory–motor binding in other not-singing tasks such as synchronized tapping to sounds (Hove et al., 2013). For instance, Schmitz et al. (2013) have shown functional connectivity between STG and basal ganglia in

an auditory–motor binding task. By presenting sounds that were congruent or not with a human avatar breaststroke, the participants were asked to judge small perceptual differences in velocity in the scanner. The STG and STS showed greater activation for sound-movement bound representations. In line with the suggested temporal coupling role, the basal ganglia and the cerebellum showed greater functional connectivity with the STG when the sounds were congruent to the movement with which they were integrated. This supports the idea that the basal ganglia and cerebellum may be crucial for sensory–motor binding to couple unfolding sound with other distinct elements. Consequently, we hypothesized that the basal ganglia and cerebellum as well as the hippocampus may play an important role in binding lyrics and melodies through covert rehearsal.

Interestingly, speech perception and memory for sentences may strongly rely on semantic and syntactic integration of words into coherent sequences (Snijders et al., 2009). These integrative processes are considered as binding and have been shown to be supported by the left IFG (Hagoort, 2005; Opitz and Friederici, 2003). This idea is in line with the dual-stream model of speech proposed by Hickok and Poeppel (2007) and is in agreement with findings reporting an encoding role in the left IFG for words (Kirchhoff et al., 2000) and semantic associative processes (Addis and McAndrews, 2006). The binding view proposed by Opitz (2010), further predicts that, although the hippocampus may be involved in the integration of the word chain into a sentence early during encoding (as a form of relational binding already described), the hierarchical rule-based integration of words within their grammatical role and position depends on the left IFG. Based on this literature, it might be possible that the left IFG could be selectively involved in binding contiguous information, at least within the verbal domain. To this regard, melody is an inherent characteristic of the lyrics in the unified condition, hence a stronger involvement of the left IFG than the hippocampus could be expected. Although the integration of linguistic content embedded in a melody has not been studied yet, activation in the left IFG has been observed during musical imitation and vocal learning in singing (Brown et al., 2004) as well as during perceptual integration of lyrics and melodies of songs (Schön et al., 2010). This evidence suggests the potential involvement of the left IFG to support the mnemonic binding of sung lyrics, particularly following the covert rehearsal imitating a given song.

In sum, the present study investigated the cerebral network involved in memory binding of songs in two different conditions, where lyrics and melodies were presented either in a unified, or in a separate way. Given the involvement of the posterior areas of the STG/STS and the left IFG in the processing of integrated song representations and the left IFG in binding verbal information and vocal imitation, we predicted that these regions would show a SME in the unified condition. Based on memory models of binding, we hypothesized that the hippocampus would be more involved in binding resulting in greater SME in the separate than in the unified condition in this region. Furthermore, the separate condition may require higher temporal and mental imagery processing demands to allow the coupling of lyrics and melodies in time. Hence, we predicted a SME in the basal ganglia and the cerebellum in the separate condition.

Methods

Participants

Twenty-two healthy native French speakers participated in this study. They were all right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), and reported to have normal hearing. All participants were non-musicians and musical experience was assessed by the Music Expertise Questionnaire (Ehrlé, 1998), indicating that participants were not music experts (mean score ± SD: 5.45 ± 2.79; max score: 27 points).