



## Cognitive and neural mechanisms underlying the mnemonic effect of songs after stroke

Vera Leo<sup>a</sup>, Aleks J. Sihvonen<sup>a,b</sup>, Tanja Linnavalli<sup>a</sup>, Mari Tervaniemi<sup>a,c</sup>, Matti Laine<sup>d</sup>, Seppo Soinila<sup>e</sup>, Teppo Särkämö<sup>a,\*</sup>

<sup>a</sup> Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Finland

<sup>b</sup> Department of Neurosciences, Faculty of Medicine, University of Helsinki, Finland

<sup>c</sup> CICERO Learning, University of Helsinki, Finland

<sup>d</sup> Department of Psychology, Åbo Akademi University, Turku, Finland

<sup>e</sup> Division of Clinical Neurosciences, Turku University Hospital, Department of Neurology, University of Turku, Finland

### ARTICLE INFO

#### Keywords:

Verbal memory  
Stroke  
Aphasia  
Singing  
Serial position effect  
Chunking

### ABSTRACT

Sung melody provides a mnemonic cue that can enhance the acquisition of novel verbal material in healthy subjects. Recent evidence suggests that also stroke patients, especially those with mild aphasia, can learn and recall novel narrative stories better when they are presented in sung than spoken format. Extending this finding, the present study explored the cognitive mechanisms underlying this effect by determining whether learning and recall of novel sung vs. spoken stories show a differential pattern of serial position effects (SPEs) and chunking effects in non-aphasic and aphasic stroke patients ( $N = 31$ ) studied 6 months post-stroke. The structural neural correlates of these effects were also explored using voxel-based morphometry (VBM) and deterministic tractography (DT) analyses of structural MRI data. Non-aphasic patients showed more stable recall with reduced SPEs in the sung than spoken task, which was coupled with greater volume and integrity (indicated by fractional anisotropy, FA) of the left arcuate fasciculus. In contrast, compared to non-aphasic patients, the aphasic patients showed a larger recency effect (better recall of the last vs. middle part of the story) and enhanced chunking (larger units of correctly recalled consecutive items) in the sung than spoken task. In aphasics, the enhanced chunking and better recall on the middle verse in the sung vs. spoken task correlated also with better ability to perceive emotional prosody in speech. Neurally, the sung > spoken recency effect in aphasic patients was coupled with greater grey matter volume in a bilateral network of temporal, frontal, and parietal regions and also greater volume of the right inferior fronto-occipital fasciculus (IFOF). These results provide novel cognitive and neurobiological insight on how a repetitive sung melody can function as a verbal mnemonic aid after stroke.

### 1. Introduction

The role of music as a learning tool in enhancing verbal memory and recall has been a focus of scientific interest and debate for decades. *Musical mnemonics* refers to the idea that music, especially singing, may act as a mnemonic or learning aid, providing a structured temporal scaffolding framework that facilitates word learning (Ferrerri and Verga, 2016). Experimental studies using songs have shown that hearing the melody of a well-known song can cue the retrieval of its lyrics (Rubin, 1977) and that lyrics are effectively paired with melody also when learning unfamiliar songs (Crowder et al., 1990; Samson and Zatorre, 1991). Also studies comparing memory for novel verbal material presented in sung or spoken formats have reported better learning and/or

delayed recall of sung vs. spoken material (Calvert and Tart, 1993; Ludke et al., 2014; McElhinney and Annett, 1996; Rainey and Larsen, 2002; Tamminen et al., 2017; Wallace, 1994), also in the context of implicit (statistical) learning (Bosseler et al., 2016; Schön et al., 2008), although some studies have found no advantage when directly comparing the two presentation forms (Racette and Peretz, 2007; Thaut et al., 2005) or when controlling for differences in presentation rate (Kilgour et al., 2000).

In addition to healthy subjects, the mnemonic effects of songs have been explored also in persons with memory impairment caused by a neurological illness. An advantage of sung material over spoken material has previously been observed on the learning and recall of word lists in multiple sclerosis (MS) patients (Thaut et al., 2014) as well as in

\* Corresponding author at: Cognitive Brain Research Unit, Department of Psychology and Logopedics, University of Helsinki, Haartmaninkatu 3, P.O. Box 21, 00014 Helsinki, Finland.

E-mail address: [teppo.sarkamo@helsinki.fi](mailto:teppo.sarkamo@helsinki.fi) (T. Särkämö).

<https://doi.org/10.1016/j.nicl.2019.101948>

Received 20 November 2018; Received in revised form 5 April 2019; Accepted 19 July 2019

Available online 05 August 2019

2213-1582/ © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

the recall of unfamiliar song lyrics or texts in Alzheimer's disease (AD) patients (Moussard et al., 2014; Palisson et al., 2015; Simmons-Stern et al., 2010) and in amnesic patients (Haslam and Cook, 2002). Aphasic stroke patients have been observed to repeat the words of familiar songs (Straube et al., 2008) as well as complete the ends of phrases of familiar songs (Kasdan and Kiran, 2018) better when they are presented in sung than spoken format, whereas no advantage of sung over spoken presentation has previously been observed in the learning of unfamiliar lyrics (Hébert et al., 2003; Racette et al., 2006; Straube et al., 2008).

Recently, using a larger sample of stroke patients ( $N = 31$ ) studied longitudinally at acute and 6-month post-stroke stage, we found that stroke patients showed better learning and delayed recall of novel, narrative stories when presented in sung vs. spoken format at the 6-month stage but not at the acute stage (Leo et al., 2018). At 6 months, the relative benefit of the sung melody on story learning was seen especially in patients with mild aphasia, suggesting that aphasic patients can benefit from the sung melody as a mnemonic aid in recalling novel verbal material. At the neural level, further results from voxel-based morphometry (VBM) and deterministic tractography (DT) analyses of structural MRI data showed that in aphasic patients the sung > spoken learning effect correlated extensively with larger grey matter volume (GMV) in left frontal areas, right temporal and limbic areas, and bilateral parietal and striatal areas as well as with higher fractional anisotropy (FA) and volume in multiple frontotemporal white matter tracts, including the uncinate fasciculus (UF), the arcuate fasciculus (AF), and the inferior fronto-occipital fasciculus (IFOF) (Särkämö and Sihvonen, 2018).

Overall, based on the previous studies on healthy subjects and clinical populations, the mnemonic effect of songs in facilitating verbal learning and memory appears to occur most often when the to-be-learned verbal material is linguistically connected (phrases or sentences) instead of isolated words and the sung melody is simple in its melodic and rhythmic structure, repetitive, and slower in tempo compared to speech. In general, the melodic and rhythmic patterns of music provide a rich structure that can potentially help in combining words and phrases, identifying line lengths and stress patterns, and adding emphasis and focus on the surface characteristics of the verbal material (e.g., phonemic structure, phrasing) (Wallace, 1994). However, few studies have explored how the mnemonic effect of singing unfolds structurally across the length of the to-be-recalled materials, and how it affects different learning- and memory-related processes.

The *serial position effect* (SPE) is a classic learning-related phenomenon whereby information occurring early and late in a sequence is remembered better than information presented in the middle of the sequence (Crowder and Greene, 2000; Deese and Kaufman, 1957). Typically, in a learning task where the number of serially presented items exceeds working memory capacity, the SPE is reflected by a U-shaped curve with an increased likelihood to recall the first items (*primacy effect*, PE) and the last items (*recency effect*, RE) compared to the ones in the middle. The classic dual-storage memory model (Atkinson and Shiffrin, 1968) attributes the RE to immediate access from a short-term storage buffer and the PE to retrieval from a long-term memory storage through working memory. Although the neural basis of SPE is still rather poorly understood, there is some evidence from neuroimaging and clinical studies of verbal memory that the PE and RE are associated with different neural networks: the PE with dorsolateral prefrontal areas and the RE with inferior parietal, temporal, and hippocampal areas (Buchsbbaum et al., 2011; Düzel et al., 1996; Innocenti et al., 2013; Spalletta et al., 2016; Staffaroni et al., 2017).

Most of the evidence for SPEs comes from studies using unconnected items, such as lists of words or numbers, with aphasic patients showing impairment especially in the PE (Ivanova et al., 2018; Jefferies et al., 2008; Ostergaard and Meudell, 1984). However, the SPE has also been found for connected speech, both in healthy subjects (Freebody and Anderson, 1986; Newhouse and Holen, 1975) and in patients with brain

injury (Hall and Bornstein, 1991) and aphasia (Brodsky et al., 2003). It is possible that the learning and recall of sung verbal material could show a different pattern of SPEs compared to spoken verbal material, but experimental evidence supporting this is still scarce and limited to healthy persons. Comparing the immediate recall of digit sequences presented in spoken and sung format to healthy subjects, Silverman (2007) found a significant interaction between serial position and condition, with less bowed U curve (indicating more stable recall performance across the list, especially for the middle items) in the sung condition. In another related study, Maylor (2002) found that the degree of familiarity of song lyrics was linked to a better recall across all verses of the song, although both PE and RE were also observed.

Grouping individual items into larger units, a process known as *chunking*, is another well-known memory mechanism that can facilitate learning and optimize performance by decreasing memory load (Bor et al., 2003; Gobet et al., 2001). Chunking-based encoding of verbal material has been linked to a wide-scale bilateral network comprising of dorsolateral prefrontal, inferior parietal, and posterior temporal regions (Bor et al., 2004). Recently, chunking training in working memory has been observed to enhance language processing in aphasia (Eom and Sung, 2016). Wallace (1994) has suggested that in a sung verbal learning task, the repetitive melodic and rhythmic structure can serve as an encoding and retrieval cue by chunking consecutive items into melodic phrases and assisting in positioning and sequencing textual units, thus decreasing the likelihood that units will be misplaced and disrupt memory for succeeding units. However, there is currently very little experimental evidence on chunking in the context of music or singing, with only two studies in healthy subjects reporting greater chunking of recalled material for words presented with background music vs. silence (Ferreri et al., 2015) and via singing vs. speaking (McElhinney and Annett, 1996).

Extending our earlier findings (Leo et al., 2018), the aim of the present study was to (i) determine whether the previously observed mnemonic benefit in the learning and delayed recall of sung vs. spoken novel narrative stories 6 months post-stroke would be related to a difference in SPE and chunking effects between the tasks and aphasic vs. non-aphasic patients, and to (ii) uncover the structural neural correlates of these effects. For this purpose, we analyzed recall performance for the first, middle, and last part of the stories as well as for the number of correctly recalled consecutive words (chunk length). Our hypothesis was that the sung melody would help in combining words and linking succeeding verses together in memory, which would result in longer chunks in the sung than spoken task. We also expected that in the sung task, recall would be more stable across the story (resulting in smaller SPE) or, optionally, more enhanced for the end part of the story (larger RE). The latter effect could emerge if the repetition of the melody across verses plays a role in building the mnemonic effect, making the last part of the story more salient in short-term storage and therefore easier to recall in the sung task. Given the close coupling between prosody and music or singing (Brown, 2017; Hausen et al., 2013; Thompson et al., 2012), we also sought to explore if the potential benefits in the sung vs. spoken task on SPE and chunking would be associated with the ability to perceive linguistic and emotional prosody. Finally, given our previous results (Leo et al., 2018) and the prior evidence for the relative preservation of vocal music (Sihvonen et al., 2017a) and the benefits of singing-based interventions in aphasia (Belin et al., 1996; Schlaug et al., 2008; van der Meulen et al., 2014; Zumbansen et al., 2014), we expected that the above-mentioned SPE and chunking effects in the sung vs. spoken task would be evident particularly in aphasic patients.

## 2. Materials and methods

### 2.1. Subjects and study design

Subjects ( $N = 31$ ) were right-handed stroke patients recruited during 2013–2016 from the Department of Clinical Neurosciences at

## Melody in the sung version of SSSRT

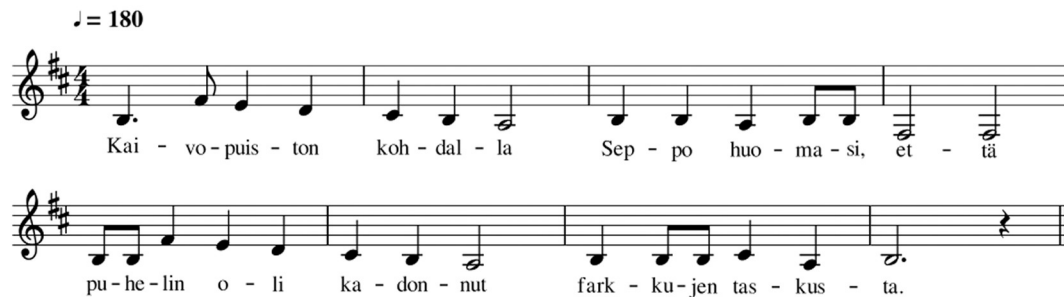


Fig. 1. Melody used in the sung part of the sung-spoken story recall task (SSSRT).

the Turku University Hospital. All patients had an MRI-verified first-ever acute ischemic stroke or intracerebral hemorrhage in the left or right hemisphere, primarily in middle cerebral artery (MCA) territory, and at least minor cognitive impairment caused by the stroke. Patients with prior neurological or psychiatric disease, substance abuse, or significant hearing impairment were excluded. All participants underwent a neuropsychological testing and a MRI session within 3 weeks after the stroke (acute stage) and at the 6-month post stroke stage. The neuropsychological testing session lasted 2–3 h. The study was approved by the Ethics Committee of the Hospital District of Southwest Finland and was performed in conformance with the Declaration of Helsinki. All patients signed an informed consent and received standard stroke treatment and rehabilitation.

### 2.2. Standard neuropsychological tests

**Aphasia** was assessed using the Aphasia Severity Rating Scale (ASRS) from the Boston Diagnostic Aphasia Examination (BDAE; Goodglass and Kaplan, 1983). The ASRS scoring was done clinically mostly based on free conversational speech, drawing information also from performance on standard tests of verbal comprehension [shortened Token Test (De Renzi and Faglioni, 1978)] and production [shortened Boston Naming Test (Laine et al., 1993a), semantic and phonemic verbal fluency tasks (Lezak et al., 2012)]. Patients with an ASRS score of 4 or less were classified as aphasic (Goodglass and Kaplan, 1983).

**Verbal memory** was evaluated with the Story Recall (SR) subtest (immediate and delayed verbal recall of a short story) from the Rivermead Behavioural Memory Test (RBMT; Wilson et al., 1985) and an Auditory Verbal Learning Task with words (AVLT; 10 orally presented words, three learning trials and delayed recall).

**Music perception** was evaluated with a shortened version (Särkämö et al., 2009) of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003) comprising of the Scale and Rhythm subtests (discriminating piano melodies based on melodic pitch and rhythm changes). The role of music in life pre-stroke was also assessed with the Barcelona Music Reward Questionnaire (BMRQ; Mas-Herrero et al., 2013).

**Prosody perception** was evaluated 6 months post-stroke with two tasks involving linguistic and emotional prosody. In the linguistic prosody task (from Hausen et al., 2013), the subjects heard 30 utterances produced with a prosodic stress pattern that denoted it either as a compound word [e.g., “näytä KISSankello” meaning “show the harebell (or literally ‘cat’s-bell’ in English) flower”] or as a phrase comprised of the same two words (e.g., “näytä KISSan KELLO” meaning “show the bell of the cat” in English). After each utterance, they saw two pictures on the screen depicting the compound and two-word phrase options and responded by selecting the picture that matched what they heard. In the emotional prosody task (adapted from Leinonen et al., 1997), the subjects heard 96 one-word utterances (female name “Saara”) produced

with a prosodic pattern that expressed six different emotional states (neutral, afraid, sad, happy, angry, surprised) and they had to select which emotion the stimulus expressed.

### 2.3. Sung-spoken story recall task

#### 2.3.1. Stimuli

The sung-spoken story recall task (SSSRT) was performed 6 months post-stroke. The task was developed to compare the learning and recall of novel verbal material (stories) presented in sung and spoken formats. The SSSRT comprised of two short narrative stories (A & B) themed around an unexpected or ironic event in everyday life (A: forgetting a mobile phone to restaurant and later discovering that someone had made expensive hotline calls with it, B: traveling to Spain on holiday and finding out that luggage was lost at the airport and arrives on the last day of the holiday). The stories were 56 (A) and 55 (B) words long and were arranged in 5 verses (10–13 words per verse). The two stories were recorded by the same female voice in (i) spoken format (with natural prosody) and (ii) in sung format. The sung versions had the same melody, which was composed to be simple, containing 6–7 different tones minor key in 4 bars, with 4/4 m and a tempo of 180 beats per minute (bpm) (see Fig. 1 for the notation of the melody). The same melody repeated in all the 5 verses of the song. The durations of stories were 34 s (A) and 36 s (B) in the spoken versions and 53 s (A & B) in the sung versions. The full lyrics and audio examples of the spoken and sung stories are available as Supplementary Material.

#### 2.3.2. Procedure

The spoken and sung versions of the SSSRT were presented to the patients by counter-balancing the verbal content of the stories. Thus, half of the patients heard story A spoken and story B sung and half heard story A sung and story B spoken. The spoken version was presented first, with three consecutive learning trials and a delayed recall trial 25 min later. Then, after a 15-min interval, the sung version was presented following the same protocol. We chose to use this fixed presentation order instead of a counterbalanced in order to avoid the possibility that when performing the sung version first, the patients could then covertly use the melody (i.e., imagining or humming it in their mind) while performing the spoken version. This would have been possible since the A and B stories were designed to have similar linguistic structure (in terms of line length and phrasing) so that both would work with the same melody. The stimuli were presented on a laptop computer with headphones, the volume being adjusted to a comfortable and clearly audible level. On each trial, the task of the patient was to try to recall as much of the story as he/she could. To make the recall situation as natural and comfortable as possible in the sung condition, the patient was given the option of recalling the story either by speaking or by singing. No patient chose singing so all recall performances were done by speaking.

**Table 1**  
Characteristics of the patients.

|   | All patients<br>N = 31 | Aphasics<br>N = 14 | Non-aphasics<br>N = 17 | p value            |
|---|------------------------|--------------------|------------------------|--------------------|
| <b>Demographical variables</b>                  |                        |                    |                        |                    |
| Age (years)                                     | 53.0 (14.3)            | 51.4 (17.7)        | 54.4 (11.3)            | 0.564 (t)          |
| Gender (male/female)                            | 19/12                  | 9/5                | 10/7                   | 0.756 ( $\chi^2$ ) |
| Education (years)                               | 14.5 (3.4)             | 15.1 (4.3)         | 14.1 (2.4)             | 0.428 (t)          |
| <b>Pre-stroke musical background</b>            |                        |                    |                        |                    |
| Formal music training (yes/no)                  | 8/23                   | 4/10               | 4/13                   | 0.750 ( $\chi^2$ ) |
| Active singing or playing (yes/no)              | 15/16                  | 7/7                | 8/9                    | 0.870 ( $\chi^2$ ) |
| BMRQ score (max.100)                            | 75.6 (12.6)            | 77.6 (11.5)        | 74.0 (13.7)            | 0.435 (t)          |
| <b>Pre-stroke leisure activities</b>            |                        |                    |                        |                    |
| Music listening frequency <sup>a</sup>          | 4.6 (1.0)              | 4.4 (1.3)          | 4.8 (0.8)              | 0.245 (t)          |
| Radio listening frequency <sup>a</sup>          | 2.7 (1.6)              | 2.4 (1.3)          | 3.0 (1.8)              | 0.328 (t)          |
| Reading frequency <sup>a</sup>                  | 3.8 (1.7)              | 3.9 (1.4)          | 3.8 (2.0)              | 0.958 (t)          |
| <b>Clinical variables (acute post-stroke)</b>   |                        |                    |                        |                    |
| Lesion laterality (left/right)                  | 20/11                  | 14/0               | 6/11                   | 0.001 ( $\chi^2$ ) |
| Lesion size (cm <sup>3</sup> )                  | 53.4 (54.5)            | 39.0 (50.4)        | 65.3 (56.2)            | 0.186 (t)          |
| Stroke type (infarct/hemorrhage)                | 22/9                   | 9/5                | 13/4                   | 0.457 ( $\chi^2$ ) |
| NIHSS score (max. 42)                           | 4.7 (3.1)              | 3.9 (2.3)          | 5.5 (3.6)              | 0.158 (t)          |
| BDAE Aphasia Severity Rating Scale <sup>b</sup> | 4.3 (0.7)              | 3.7 (0.5)          | 4.8 (0.4)              | < 0.001(t)         |
| MBEA Scale and Rhythm avg. (%)                  | 72.9 (14.1)            | 73.9 (9.9)         | 72.1 (17.1)            | 0.720 (t)          |
| AVLT Learning score (3 trials, max. 30)         | 18.0 (4.6)             | 15.6 (4.0)         | 20.0 (4.2)             | 0.006 (t)          |
| AVLT Delayed recall score (max. 10)             | 4.5 (2.7)              | 3.4 (2.4)          | 5.4 (2.7)              | 0.049 (t)          |
| RBMT Story recall immediate (max. 42)           | 13.9 (7.0)             | 12.0 (6.0)         | 15.5 (7.6)             | 0.174 (t)          |
| RBMT Story recall delayed (max. 42)             | 11.1 (7.5)             | 8.6 (5.3)          | 13.1 (8.6)             | 0.104 (t)          |

Data are reported as mean (SD) unless otherwise stated. Abbreviations: t = independent-samples t-test,  $\chi^2$  = chi-square test, AVLT = Auditory Verbal Learning Task, BDAE = Boston Diagnostic Aphasia Examination, BMRQ = Barcelona Music Reward Questionnaire, MBEA = Montreal Battery of Evaluation of Amusia, NIHSS = National Institute of Health Stroke Scale, RBMT = Rivermead Behavioural Memory Test.

<sup>a</sup> Likert scale 1–7 (1 = not at all, 7 = daily).

<sup>b</sup> Likert scale 0–5 (0 = no usable speech or comprehension, 5 = minimal speech handicaps).

### 2.3.3. Data analysis

The scoring protocol for the SSSRT was similar as in the RBMT-SR: 2 points for each correct word and 1 point for each partially correct or semantically similar word. For the SPE analyses, the percentages of correct responses for the three learning trials (T1/T2/T3) and the delayed recall (del) trial were calculated separately for the first verse (V1), the middle verse (V3), and the last verse (V5), as well as for the primacy effect (PE, V1 minus V3) and the recency effect (RE, V5 minus V3). For the chunking analyses, we first divided the item-level (individual words) scores across the stories into chunks (defined as the number of consecutive words that were recalled correctly). Based on this, we then calculated the average length of the chunks in the two tasks. Differences between the sung and spoken task performance were analyzed statistically using mixed-model analyses of variance (ANOVA) as well as independent-samples and paired *t*-tests. In the mixed-model ANOVAs, the Greenhouse-Geisser correction was used when appropriate. The level of statistical significance was set at  $p < .05$ . All statistical analyses were performed using IBM SPSS Statistics 24.

## 2.4. Structural magnetic resonance imaging

### 2.4.1. MRI data acquisition

Participants were scanned 6 months post-stroke using a standard 12-channel head matrix coil on a 3 T Siemens Magnetom Verio scanner (Siemens Medical Solutions, Erlangen, Germany) at the Medical Imaging Centre of Southwest Finland. T1-weighted high-resolution MPRAGE scans (flip angle = 9°, TR = 2300 ms, TE = 2.98 ms, voxel size = 1.0 × 1.0 × 1.0 mm) as well as diffusion MRI scans (TR = 11,700 ms, TE = 88 ms, acquisition matrix = 112 × 112, 66 axial slices, voxel size = 2.0 × 2.0 × 2.0 mm) with one non-diffusion weighted volume and 64 diffusion weighted volumes (b-values of 1000 s/mm<sup>2</sup>) were acquired.

### 2.4.2. Voxel-based morphometry

Voxel-based morphometry (VBM) analysis was carried out using

VBM (Ashburner and Friston, 2000) Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology, UCL) under MATLAB 8.4.0 (The MathWorks Inc., Natick, MA, USA, version R2014b). For exact methodological description including detailed information on preprocessing, see our recently published articles (Sihvonen et al., 2016; Sihvonen et al., 2017b). The preprocessed and modulated grey matter (GM) images were entered into a second-level analysis using *t*-tests to assess the relationship between the behavioural performance and the GM volume (GMV) across the entire GM space within the aphasic group and in the aphasic vs. non-aphasic group. Age, gender, and total intracranial volume (TIV) were added as nuisance covariates (Barnes et al., 2010). All results were thresholded at a whole-brain uncorrected  $p < .001$  threshold with a cluster extent of > 100 contiguous voxels. To evaluate which GM correlates were facilitating the behavioural performance, partial correlations with two-tailed false discovery rate (FDR) corrected *p*-values controlling for age, sex and TIV were calculated for each significant cluster separately for aphasic and non-aphasic patients.

### 2.4.3. Deterministic tractography

To evaluate the relationship between the task performance and white matter (WM) pathways, the following four tracts were dissected in the left and right hemispheres using TrackVis (version 0.6.0.1, Build 2015.04.07) and included in the deterministic tractography (DT) analyses: arcuate fasciculus (AF), inferior fronto-occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF), and uncinate fasciculus (UF). These frontotemporal tracts were selected on the basis of previous DTI studies in healthy subjects and different clinical populations that link them directly to verbal learning or verbal memory performance (Chiou et al., 2016; López-Barroso et al., 2013; Mabbott et al., 2009; McDonald et al., 2008; Reggente et al., 2018). For complete methodological information including description on the included WM tracts, see our recently published article (Sihvonen et al., 2017c). After dissection, statistical information (tract volume and FA value) of each WM tract was collected using a MATLAB toolbox, “along-tract statistics” (Colby



et al., 2012). Volume and FA values were then imported to IBM SPSS Statistics 24 and further analyzed to evaluate the relationship between the WM tract parameters and behavioural performance using two-tailed Pearson correlation analysis in the aphasic and non-aphasic groups. Standard FDR-correction was applied to control for multiple correlations.

### 3. Results

#### 3.1. Patient characteristics

Based on the BDAE-ASRS scores, 14 patients (45%) were classified as aphasic (all with left hemisphere lesions) and 17 as non-aphasic. In the aphasic group, the severity of language impairment was primarily mild (10 patients: BDAE-ASRS score 4, 4 patients: BDAE-ASRS score 3). As shown in Table 1, the aphasic and non-aphasic groups were comparable in all demographic and clinical characteristics, except in AVLT learning [ $t(29) = 2.9$ ,  $p = .006$ ] and delayed recall [ $t(29) = 2.1$ ,  $p = .049$ ] in which the aphasic patients scored lower, as expected.

#### 3.2. Serial position effects in the sung vs. spoken tasks

Verse-level performance of the patients in the SSSRT is shown in Table 2. Differences between SPEs in the sung and spoken tasks were analyzed using a mixed-model ANOVA with Task (spoken/sung), Trial (T1/T2/T3/del), and Verse (V1/V3/V5) as within-subject factors and Aphasia (non-aphasic/aphasic) as a between-subject factor. To focus here on SPE, we report only the main and interaction effects involving

**Table 2**  
Verse-level performance in the SSSRT.

| Task             | Trial                      | Verse | All patients | Aphasic     | Non-aphasic |
|------------------|----------------------------|-------|--------------|-------------|-------------|
|                  |                            |       | N = 31       | N = 14      | N = 17      |
| Spoken           | 1st learning trial (T1)    | V1    | 57.7 (23.0)  | 54.8 (20.4) | 60.2 (25.2) |
|                  |                            | V3    | 35.2 (27.7)  | 39.3 (31.1) | 31.8 (25.0) |
|                  |                            | V5    | 44.4 (28.5)  | 38.3 (32.2) | 49.5 (25.0) |
|                  | 2nd learning trial (T2)    | V1    | 67.6 (24.9)  | 64.7 (28.6) | 70.0 (22.1) |
|                  |                            | V3    | 45.7 (23.7)  | 45.4 (22.9) | 45.8 (25.0) |
|                  |                            | V5    | 61.9 (23.1)  | 58.8 (23.1) | 64.4 (23.4) |
|                  | 3rd learning trial (T3)    | V1    | 72.3 (25.1)  | 65.4 (31.4) | 78.1 (17.4) |
|                  |                            | V3    | 58.1 (22.4)  | 56.3 (21.1) | 59.5 (24.0) |
|                  |                            | V5    | 62.3 (28.0)  | 58.1 (29.1) | 65.8 (27.5) |
|                  | Delayed recall trial (del) | V1    | 56.5 (27.0)  | 58.3 (31.1) | 54.9 (23.9) |
|                  |                            | V3    | 54.2 (21.4)  | 55.6 (21.0) | 53.0 (22.3) |
|                  |                            | V5    | 53.2 (25.6)  | 47.4 (30.1) | 58.0 (20.9) |
|                  | Average across trials      | V1    | 63.5 (22.4)  | 60.8 (25.3) | 65.8 (20.2) |
|                  |                            | V3    | 48.3 (19.0)  | 49.2 (18.0) | 47.5 (20.3) |
|                  |                            | V5    | 55.5 (21.9)  | 50.6 (23.3) | 59.4 (20.5) |
| PE (V1 minus V3) |                            |       | 15.2 (20.3)  | 11.6 (14.8) | 18.2 (23.9) |
| RE (V5 minus V3) |                            |       | 7.2 (16.6)   | 1.5 (14.7)  | 11.9 (17.0) |
| Sung             | 1st learning trial (T1)    | V1    | 49.3 (26.5)  | 51.4 (33.0) | 47.8 (20.6) |
|                  |                            | V3    | 34.9 (24.8)  | 33.7 (21.5) | 35.9 (27.8) |
|                  |                            | V5    | 50.3 (22.1)  | 51.0 (25.0) | 49.7 (20.2) |
|                  | 2nd learning trial (T2)    | V1    | 69.8 (22.4)  | 76.9 (20.2) | 64.0 (23.0) |
|                  |                            | V3    | 60.9 (24.2)  | 60.9 (31.3) | 60.8 (17.4) |
|                  |                            | V5    | 65.2 (19.0)  | 67.2 (21.3) | 63.6 (17.5) |
|                  | 3rd learning trial (T3)    | V1    | 71.4 (24.8)  | 73.5 (27.6) | 69.5 (22.7) |
|                  |                            | V3    | 66.6 (23.9)  | 64.5 (26.4) | 68.3 (22.1) |
|                  |                            | V5    | 68.3 (19.5)  | 73.1 (21.5) | 64.2 (17.2) |
|                  | Delayed recall trial (del) | V1    | 62.5 (25.3)  | 64.5 (29.6) | 60.9 (22.0) |
|                  |                            | V3    | 62.5 (26.1)  | 54.6 (31.7) | 69.0 (19.0) |
|                  |                            | V5    | 63.9 (18.5)  | 66.6 (20.3) | 61.8 (17.2) |
|                  | Average across trials      | V1    | 63.9 (21.7)  | 66.6 (24.9) | 61.6 (19.1) |
|                  |                            | V3    | 56.7 (20.3)  | 53.5 (23.9) | 59.5 (16.9) |
|                  |                            | V5    | 62.2 (16.9)  | 64.4 (19.6) | 60.2 (14.4) |
| PE (V1 minus V3) |                            |       | 7.3 (19.7)   | 13.1 (15.6) | 2.1 (21.9)  |
| RE (V5 minus V3) |                            |       | 5.5 (17.2)   | 11.0 (15.4) | 0.8 (17.6)  |

Data are mean (SD). PE = primacy effect, RE = recency effect, V1 = 1st verse, V3 = 3rd (middle) verse, V5 = 5th (last) verse.

Verse. The mixed-model ANOVA yielded a significant Verse main effect [ $F(2, 56) = 17.2$ ,  $p < .001$ ], indicating better recall of V1 than V3 (primacy effect, PE) and V5 than V3 (recency effect, RE). We also observed a Trial x Verse interaction [ $F(3.6, 99.5) = 4.4$ ,  $p = .004$ ], indicating more uniform recall of the verses in the delayed recall trial than in the learning trials, and Task x Verse x Aphasia interaction [ $F(1.5, 40.9) = 3.7$ ,  $p = .046$ ] (see Fig. 2).

This latter three-way interaction was followed up by separate mixed-model ANOVAs performed for the non-aphasic and aphasic patients and for the spoken and sung tasks. In the non-aphasic patients, the Verse effect was significant in the spoken task [ $F(2, 32) = 7.9$ ,  $p = .002$ ], with post hoc tests showing better recall of V1 than V3 (PE;  $p = .006$ ) and V5 than V3 (RE;  $p = .011$ ), but not in the sung task (Fig. 2). In the aphasic patients, the Verse effect was significant in both the spoken task [ $F(2, 26) = 5.5$ ,  $p = .010$ ] and the sung task [ $F(2, 26) = 6.3$ ,  $p = .006$ ]. Post hoc tests showed that in the spoken task the aphasic patients recalled V1 better than V3 (PE;  $p = .012$ ) and also V1 better than V5 ( $p = .014$ ), whereas in the sung task they recalled V1 better than V3 (PE;  $p = .008$ ) and V5 better than V3 (RE;  $p = .019$ ) (Fig. 2).

The differences in the SPE patterns of the tasks were further analyzed with (i) paired *t*-tests comparing the sung and spoken tasks for each verse (V1/V3/V5) and for the PE (V1 minus V3) and RE (V5 minus V3) within the non-aphasic and aphasic patient groups and (ii) independent-samples *t*-tests comparing the non-aphasic and aphasic patient groups on the difference between the tasks (sung minus spoken) for each verse (V1/V3/V5), PE, and RE. The paired *t*-tests showed that in the non-aphasic patients, recall was better in the sung than spoken task for V3 [ $t(15) = 2.7$ ,  $p = .018$ ] (Fig. 2) and both the PE and the RE were marginally smaller in the sung than spoken task [ $t(15) = 2.0$ ,  $p = .063$  and  $t(15) = 2.1$ ,  $p = .055$ , respectively]. In contrast, the aphasic patients showed better recall of V5 in the sung than spoken task [ $t(13) = 3.3$ ,  $p = .006$ ] (Fig. 2). The independent-samples *t*-tests indicated that the recall of V5 was better in the sung than spoken task in the aphasic than in the non-aphasic patients [ $t(28) = 2.5$ ,  $p = .019$ ]. Similarly, the RE was larger in the sung than spoken task in the aphasic than in the non-aphasic patients [ $t(28) = 2.6$ ,  $p = .014$ ].

Together, this pattern of results suggests that in non-aphasic patients the sung presentation led to more stable recall across the story, reducing the SPE and facilitating the recall of the middle part compared to the spoken presentation. In contrast, in aphasic patients the sung presentation facilitated the recall of the last part of the story, resulting in larger RE, compared to the spoken presentation.

#### 3.3. Chunking effects in the sung vs. spoken tasks

Chunk length averages of the patients in the SSSRT are presented in Table 3. Differences between chunk length in the sung and spoken tasks were analyzed using a mixed-model ANOVA with Task (spoken/sung) and Trial (T1/T2/T3/del) as within-subject factors and Aphasia (non-aphasic/aphasic) as a between-subject factor. This yielded a significant Trial main effect [ $F(1.8, 53.1) = 15.8$ ,  $p < .001$ ], indicating progressively increasing chunk length during the learning trials, as well as a significant Task x Aphasia interaction [ $F(1, 29) = 7.4$ ,  $p = .011$ ] (see Fig. 3).

The Task x Aphasia interaction was further analyzed with (i) paired *t*-tests comparing the chunk length between the sung and spoken tasks within the non-aphasic and aphasic groups and (ii) an independent-samples *t*-test comparing the non-aphasic and aphasic groups on the chunk length difference between the tasks (sung minus spoken). The paired *t*-tests showed an opposite, marginally significant pattern of recall effects in the two tasks, with the aphasic patients recalling longer chunks in the sung than spoken task [ $t(13) = 2.0$ ,  $p = .073$ ] and the non-aphasic patients recalling longer chunks in the spoken than sung task [ $t(15) = 1.8$ ,  $p = .097$ ]. The independent-samples *t*-test showed that the aphasic patients recalled longer chunks in the sung than spoken

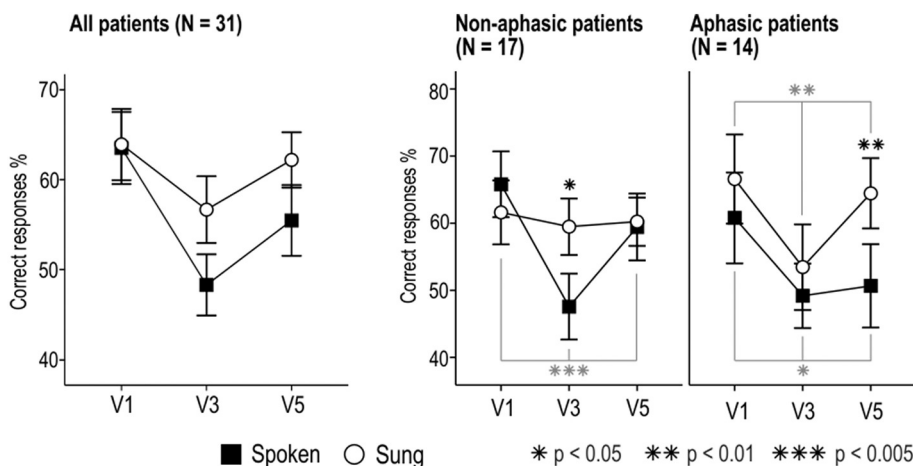


Fig. 2. Percentage of correct responses (mean ± SEM) in the first (V1), middle (V3) and last (V5) verses of the sung (white circles) and spoken (black squares) story recall tasks. Data are shown across all patients (left) and within the non-aphasic and aphasic groups (right). Grey asterisks denote significant Verse main effects in mixed-model ANOVAs in the sung and spoken tasks. Black asterisks denote significant differences between the sung and spoken tasks for individual verses in paired t-tests.

Table 3  
Average chunk length in the SSSRT.

| Task   | Trial                      | All patients | Aphasic   | Non-aphasic |
|--------|----------------------------|--------------|-----------|-------------|
|        |                            | N = 31       | N = 14    | N = 17      |
| Spoken | 1st learning trial (T1)    | 3.1 (1.5)    | 3.1 (1.6) | 3.1 (1.4)   |
|        | 2nd learning trial (T2)    | 3.7 (1.6)    | 3.5 (1.8) | 3.8 (1.5)   |
|        | 3rd learning trial (T3)    | 4.9 (3.0)    | 4.3 (2.1) | 5.5 (3.5)   |
|        | Delayed recall trial (del) | 3.8 (2.5)    | 3.3 (1.9) | 4.2 (2.8)   |
|        | Average across trials      | 3.9 (1.9)    | 3.6 (1.6) | 4.2 (2.1)   |
| Sung   | 1st learning trial (T1)    | 2.8 (1.2)    | 3.1 (1.5) | 2.5 (0.8)   |
|        | 2nd learning trial (T2)    | 3.3 (1.6)    | 3.5 (2.1) | 3.1 (0.9)   |
|        | 3rd learning trial (T3)    | 5.6 (4.5)    | 6.5 (6.2) | 4.7 (2.0)   |
|        | Delayed recall trial (del) | 4.2 (2.7)    | 4.7 (3.6) | 3.8 (1.5)   |
|        | Average across trials      | 4.0 (2.1)    | 4.4 (2.8) | 3.6 (1.1)   |

Data are mean (SD).

task compared to the non-aphasic patients [t(28) = 2.7, p = .013].

3.4. Relationship between prosody perception and SSSRT serial position and chunking effects

In order to determine if the mnemonic benefit of sung compared to spoken presentation was associated with prosodic skills, we performed

correlation analyses (Pearson, two-tailed) between the scores on the linguistic and emotional prosody perception tasks and performance in the spoken and sung tasks for each verse (V1, V3, V5), PE, RE, and chunk length in non-aphasic and aphasic patients. There were no significant correlations in non-aphasic patients. In aphasic patients (see Fig. 4), emotional prosody perception correlated significantly with the recall of V1 (r = 0.53, p = .05) and V5 (r = 0.65, p = .011) in the spoken task and of V1 (r = 0.77, p = .001), V3 (r = 0.76, p = .002), and V5 (r = 0.86, p < .001) in the sung task. Analyses of the difference score (sung minus spoken) further showed that the better recall of V3 in the sung than spoken task correlated with better emotional prosody perception (r = 0.58, p = .031). In aphasic patients, emotional prosody perception also correlated with chunk length in the sung task (r = 0.76, p = .001) and in the difference between the tasks (sung minus spoken, r = 0.79, p = .001). No other significant correlations were observed within the aphasic group.

3.5. Neural correlates of the serial position and chunking effects in the sung vs. spoken tasks

The structural neural correlates of the different SPE patterns and chunking effects in the sung vs. spoken task in the aphasic and non-aphasic patients (see above) were analyzed by correlating them with grey matter volume (GMV) from voxel-based morphometry (VBM) data

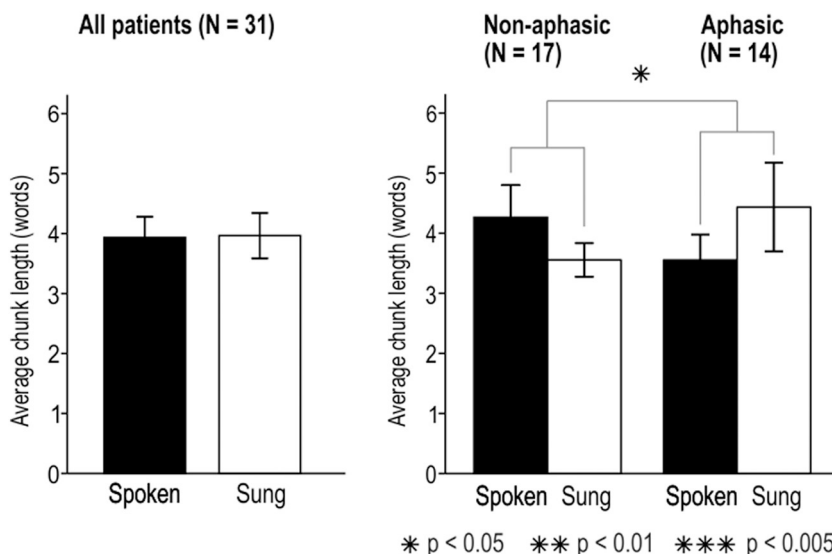
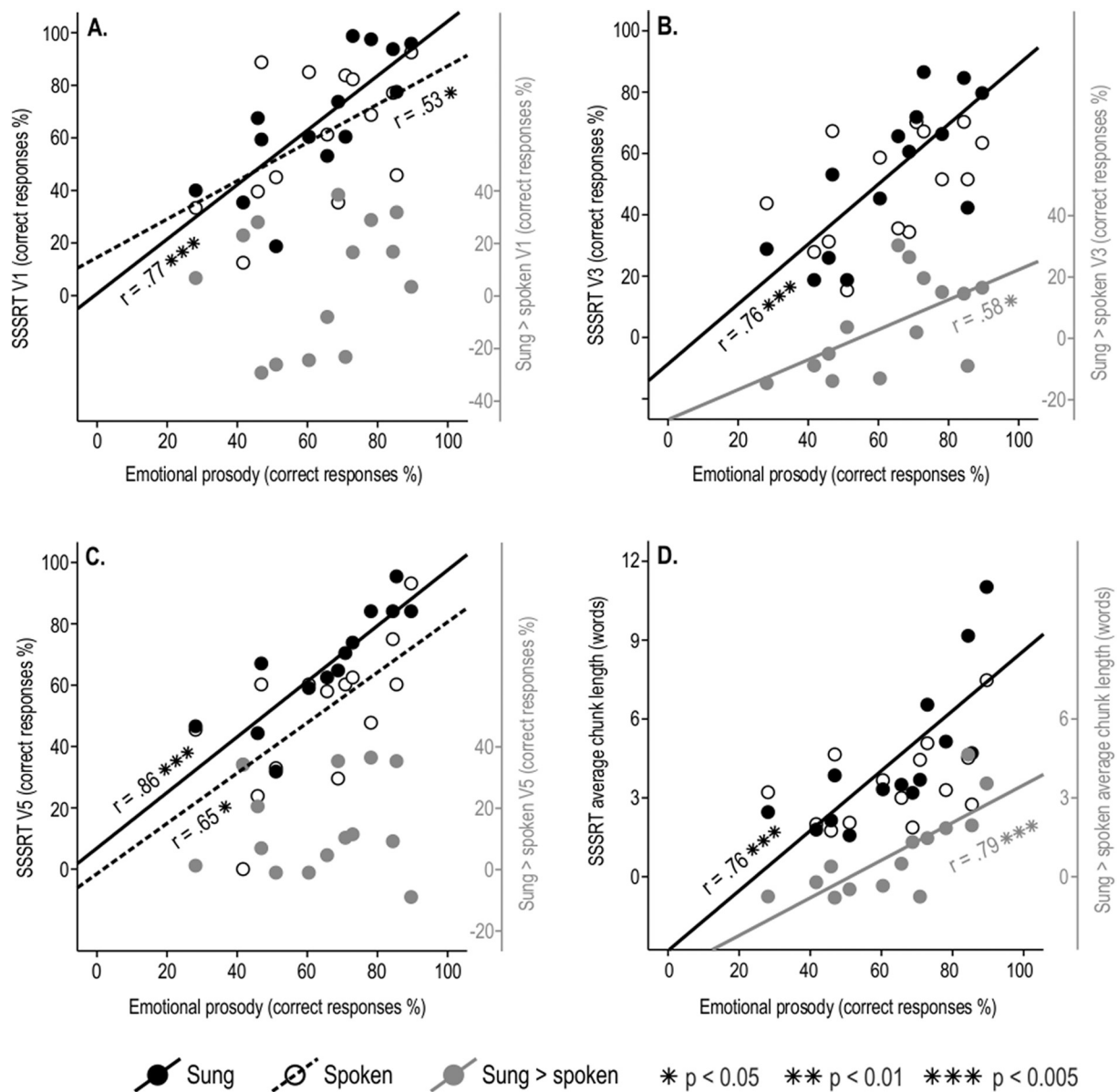


Fig. 3. Average length of recalled chunks (mean ± SEM) in the spoken (black) and sung (white) tasks in all patients (left) and in non-aphasic and aphasic patients (right). Significant Task x Group interaction indicated with an asterisk.



**Fig. 4.** Scatter plots showing the correlation between the recall of V1 (A), V3 (B), and V5 (C) and average chunk length (D) in the sung (black) and spoken (white) tasks and in their difference (sung minus spoken, grey) and the performance in the emotional prosody perception task in aphasic patients ( $N = 14$ ). Only significant correlations are shown with regression lines.

and to the volume and fractional anisotropy (FA) of white matter tracts from deterministic tractography (DT) data.

### 3.5.1. VBM results

In the aphasic patients, the larger RE in the sung vs. spoken task was associated with greater GMV in the left posterior temporal [superior temporal gyrus (STG), middle temporal gyrus (MTG)], parietal [post-central gyrus (postCG), middle occipital gyrus (MOG)], and limbic [parahippocampal gyrus (PHG)] regions as well as in the right frontal [precentral gyrus (preCG)], posterior temporal [inferior temporal gyrus (ITG)], and parietal [inferior parietal lobule (IPL)] regions (Table 4 and Fig. 5A). Similarly, the left posterior temporal (MTG), frontal (preCG), and parietal (postCG, MOG) regions as well as the right posterior temporal (ITG) regions also showed greater GMV associated with the sung > spoken RE in aphasic compared to non-aphasic patients (Table 4 and Fig. 5B). No other significant effects were observed.

### 3.5.2. DT results

In the aphasic patients (Fig. 6A), there was a significant correlation between larger RE in the sung than spoken task and larger volume of the right inferior fronto-occipital fasciculus (IFOF;  $r = 0.67$ ,  $p = .009$ ). In contrast, the non-aphasic patients (Fig. 6B) showed a strong correlation between smaller PE in the sung than spoken task and larger volume ( $r = -0.69$ ,  $p = .003$ ) and FA ( $r = -0.63$ ,  $p = .009$ ) of the left arcuate fasciculus (AF, long segment). No other significant correlations were observed.

## 4. Discussion

In the present study, we set out to investigate in a sample of 31 stroke patients whether the mnemonic benefit of sung vs. spoken novel narrative stories after stroke (Leo et al., 2018) would be related to a difference in the memory processes involved, indicated by serial position and chunking effects, and whether these effects would differ in

**Table 4**  
Significant correlations between grey matter volume and SSSRT performance.

| Patients/Contrast     | Condition        | Area                                       | MNI coordinates | Cluster size | t-Value | Correlation          |
|-----------------------|------------------|--|-----------------|--------------|---------|----------------------|
| Aphasic               | Sung > spoken RE | Left middle temporal gyrus (BA 21)         | -66 -52 1       | 1008         | 15.2**  | $r = 0.96, p < .001$ |
|                       |                  | Left superior temporal gyrus (BA 22)       | -69 -41 11      |              |         |                      |
|                       |                  | Left middle occipital gyrus (BA 19)        | -45 -88 13      | 512          | 14.5*   | $r = 0.95, p < .001$ |
|                       |                  | Left postcentral gyrus (BA 3)              | -35 -21 44      | 849          | 9.4**   | $r = 0.95, p < .001$ |
|                       |                  | Left parahippocampal gyrus (BA 34)         | -13 -19 -22     | 643          | 8.2*    | $r = 0.95, p < .001$ |
|                       |                  | Right inferior temporal gyrus (BA 20, 37)  | 47 -79 -24      | 2596         | 8.5**   | $r = 0.96, p < .001$ |
|                       |                  | Right cerebellum                           | 53 -65 -25      |              |         |                      |
|                       |                  | Right inferior parietal lobule (BA 40)     | 59 -31 34       | 1184         | 8.1**   | $r = 0.93, p < .001$ |
|                       |                  | Right postcentral gyrus (BA 2, 3)          | 57 -22 45       |              |         |                      |
|                       |                  | Right precentral gyrus (BA 4)              | 57 -14 36       | 480          | 7.3*    | $r = 0.95, p < .001$ |
| Aphasic > Non-aphasic | Sung > spoken RE | Left middle temporal gyrus (BA 37)         | -46 -82 15      | 880          | 5.7*    | n.s.                 |
|                       |                  | Left middle occipital gyrus (BA 19)        | -50 -75 0       |              |         |                      |
|                       |                  | Left middle temporal gyrus (BA 21, 37, 39) | -61 -56 5       | 782          | 5.6*    | n.s.                 |
|                       |                  | Left precentral gyrus (BA 4)               | -41 -23 57      | 1026         | 5.0*    | $r = 0.74, p = .003$ |
|                       |                  | Left postcentral gyrus (BA 3)              | -39 -26 57      |              |         |                      |
|                       |                  | Right inferior temporal gyrus (BA 20, 37)  | 59 -56 -24      | 1874         | 6.0**   | $r = 0.55, p = .041$ |
|                       |                  | Right cerebellum                           | 47 -70 -23      |              |         |                      |

All results are thresholded at a whole-brain uncorrected  $p < .001$  threshold at the voxel level with a minimal cluster size set to 100 voxels. Correlations are partial correlations with 2-tailed  $p$ -value controlling for age, sex and TIV. BA = Brodmann area, RE = recency effect.

\*  $p < .05$  FWE-corrected at the cluster level.

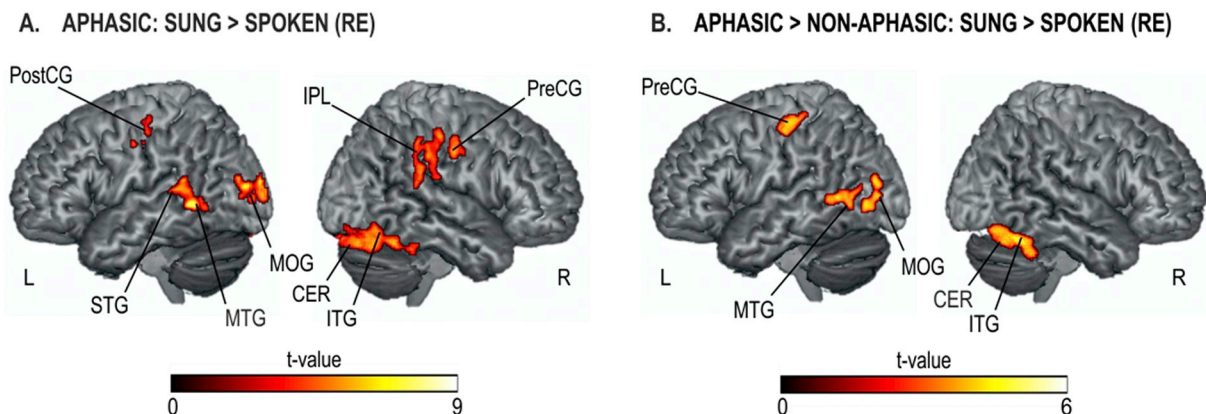
\*\*  $p < .005$  FWE-corrected at the cluster level.

aphasic and non-aphasic patients. Moreover, using voxel-based morphometry (VBM) and deterministic tractography (DT) analyses, we sought to uncover the structural neural correlates of these effects. Our main finding was that non-aphasic and aphasic patients showed a different pattern of serial position effects (SPE) and chunking effects in the two tasks, with (i) more stable recall performance (no SPE) in the sung than spoken task in the non-aphasic patients and (ii) longer recalled chunks and better recall of last verse (larger recency effect, RE) in the sung than spoken task in the aphasic patients. The latter effect was also coupled with grey matter volume (GMV) in mostly bilateral temporal, frontal, and parietal regions, as well as with the volume of the right inferior fronto-occipital fasciculus (IFOF).

In the non-aphasic patients, there was a classic SPE pattern, with both primacy effect (PE) and RE in the spoken task, but no discernible SPE in the sung task. Verse-level comparison between the tasks showed that non-aphasic patients recalled the middle verse (V3) better in the sung than spoken task, thereby making the recall performance more stable in the sung task. This result is in line with previous studies in healthy subjects, which have reported a smaller SPE (less bowed U curve) for the learning of digits when presented in sung than spoken format (Silverman, 2007) and a familiarity effect whereby familiar song

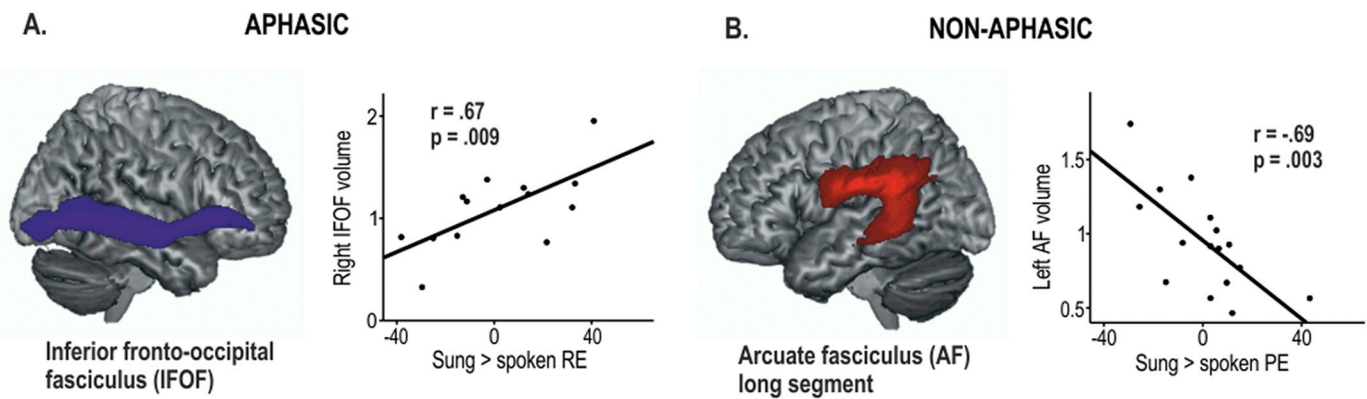
lyrics are linked to better recall performance across all verses of a song (Maylor, 2002). Interestingly, the better stability of recall in the sung task, which resulted in a smaller PE than in the spoken task, also correlated with the volume and integrity (indicated by fractional anisotropy) of the left arcuate fasciculus (AF) in non-aphasic patients. Forming the dorsal “perception-action” pathway, the left AF is thought to map sensory targets in posterior temporal areas to motor programs coded in Broca’s area (Hickok and Poeppel, 2007) and it has also been found to be crucial for sustained rehearsal processes in verbal working memory (Buchsbbaum et al., 2005). Consequently, the left AF is considered an important tract for verbal learning in children (Leroy et al., 2011; Su et al., 2018) and adults (López-Barroso et al., 2013; Thiebaut de Schotten et al., 2014) and it also forms a key part of the dual-stream pathway for singing production (Loui, 2015), its structure being malleable by singing training (Halwani et al., 2011). It is plausible that the slower presentation rate and the cues provided by the repetitive melodic structure in our sung task made it easier for covert rehearsal and less demanding for working memory, thereby recruiting the left dorsal pathway (AF) less than the spoken task.

Aphasic patients showed a PE in both the spoken and the sung task but a RE only in the sung task. The RE was larger in the sung than



**Fig. 5.** Voxel-based morphometry (VBM) results showing significant correlations between regional grey matter volume and the sung > spoken recency effect (RE) in the left and right hemispheres (A) within the aphasic group and (B) in a contrast between the aphasic and non-aphasic groups. CER = cerebellum, IPL = inferior parietal lobule, ITG = inferior temporal gyrus, L = left, MOG = middle occipital gyrus, MTG = middle temporal gyrus; PostCG = postcentral gyrus, PreCG = precentral gyrus, R = right, STG = superior temporal gyrus.





**Fig. 6.** Deterministic tractography (DT) results showing significant correlations between (A) the volume of the right inferior fronto-occipital fasciculus (IFOF) and the sung > spoken recency effect (RE) in aphasic patients and (B) the volume of the left arcuate fasciculus (AF, long segment) and the sung > spoken primacy effect (PE) in non-aphasic patients.

spoken task in aphasic compared to non-aphasic patients, which was attributable primarily to aphasic patients' better recall of the last verse (V5) in the sung task. Aphasic patients also recalled longer chunks in the sung than spoken task compared to non-aphasic patients. This pattern of results is line with the hypothesis that (i) the sung melody helps in combining words and linking succeeding verses together in memory, enabling chunking (Ferreri et al., 2015; McElhinney and Annett, 1996; Wallace, 1994) and (ii) the repetition of the same melody across the verses gradually builds the mnemonic effect in the sung task, making the last part of the story more salient and easily accessible for recall (enhancing the RE) once the melody is learned. Previous studies in non-fluent aphasic patients have shown that when recalling sequential verbal material, aphasic patients have a relatively normal RE but a clearly smaller PE, which seems to be related to their reduced verbal memory span and difficulties in covert rehearsal (Ivanova et al., 2018; Jefferies et al., 2008; Ostergaard and Meudell, 1984). Thus, it is possible that due to their verbal memory deficits, the aphasic patients in our study did not benefit from the sung melody as an aid in covert rehearsal in working memory as the non-aphasic patients (who had less severe memory deficits) apparently did, but showed more automatic effects driven more by stimulus-specific factors (i.e., the repetitive melody and rhythm of the song enable chunking and make the last verse most salient in short-term storage). It is possible that recalling the sung melody may also act as a contextual cue when attempting to retrieve the lyrics from memory. For example, Isarida and Isarida (2006) found that in the delayed recall of word lists, the RE was stronger when same environmental cues (including also background music played during the task) were present in both the initial encoding and the delayed recall situation in healthy subjects.

In VBM, the sung > spoken RE was associated with greater GMV (i) in left posterior temporal (STG, MTG), parietal (postCG, MOG), and limbic (PHG) regions and in right frontal (preCG), posterior temporal (ITG), and parietal (IPL) regions within the aphasic group and (ii) in left posterior temporal (MTG), frontal (preCG), and parietal (postCG, MOG) and right posterior temporal (ITG) regions in the aphasic group compared to the non-aphasic group. This pattern was different than the one we previously reported for the general sung > spoken learning effect in aphasic patients, which comprised of primarily left prefrontal areas [inferior frontal gyrus (IFG), middle frontal gyrus (MFG), anterior cingulate] and bilateral superior parietal and striatal areas (Särkämö and Sihvonen, 2018). The present results are well in line with previous neuroimaging studies of the RE in verbal memory in both healthy subjects and clinical groups with memory deficits, which have specifically implicated the IPL (Buchsbaum et al., 2011; Innocenti et al., 2013), MTG and ITG (Düzel et al., 1996; Spalletta et al., 2016; Staffaroni et al., 2017), and hippocampal area (Spalletta et al., 2016) in the RE. Overall, the results also conform well with neuroimaging studies of singing in which the STG and MTG have been linked to

perceptual processing of lexical/phonological and melodic features and the preCG, postCG, and cerebellar regions to vocal-motor processing, in a largely bilateral fashion (Callan et al., 2006; Méndez Orellana et al., 2014; Özdemir et al., 2006; Salmi et al., 2017; Schön et al., 2010; Segado et al., 2018). Aside from production, also listening to singing activates bilateral frontotemporal areas as well as subcortical/limbic areas (e.g., hippocampus, striatum, orbitofrontal cortex) more extensively than listening to speech (Callan et al., 2006; Méndez Orellana et al., 2014; Schön et al., 2010). The right IPL, in turn, has been identified as a crucial hub for higher-level analysis of melodic (Royal et al., 2016) and rhythmic (Konoike et al., 2012) structure of music.

In neuroimaging studies of aphasic patients, recovery of speech has been linked to the functioning of both left and right hemisphere language networks (Forkel et al., 2014; Saur et al., 2006) and singing-based rehabilitation, for example using melodic intonation therapy (MIT), has been found to increase functional activation in fronto-temporal auditory-motor and language areas during speech/singing production, either in the left or right hemisphere or bilaterally (Belin et al., 1996; Breier et al., 2010; Jungblut et al., 2014; Laine et al., 1993b; Schlaug et al., 2008). Using DTI, structural neuroplasticity changes induced by MIT have thus far been reported only in right frontotemporal tracts, including the AF and the uncinate fasciculus (UF) (Wan et al., 2014; Zipse et al., 2012). In the present study, we found that the sung > spoken RE in aphasic patients correlated with larger volume in the right IFOF. Part of the ventral processing stream, the IFOF has been increasingly recognized to play a role in language processing (Dick and Tremblay, 2012) and cognition (Cremers et al., 2016), including working memory and learning (Chiou et al., 2016; Krogsrud et al., 2018), as well as in music perception (Sihvonen et al., 2017c). In addition to occipital lobe areas, the posterior termination branches of the IFOF extend also to the posterior inferior temporal and parietal areas (Duffau et al., 2013; Sarubbo et al., 2013), close to the ITG and IPL clusters that were linked to the RE in our aphasic patients and in previous imaging studies (see above). As the IFOF is thought to play a role in multimodal integration and semantic processing, (Sarubbo et al., 2013), it is plausible that in our sung task this tract carries the melodic and rhythm information to prefrontal areas (IFG, MFG) where this is integrated with the verbal content of the story and processed in working memory.

In addition to music, the right ventral stream has recently been implicated in the perception of prosody (Frühholz et al., 2015; Sammler et al., 2015). Interestingly, the correlation analyses showed that in aphasic patients, performance in the emotional (but not linguistic) prosody perception task was linked to better recall of the first (V1) and last (V5) verses in the spoken task and more strongly with better recall of the first (V1), middle (V3), and last (V5) verses in the sung task, with the sung > spoken difference and emotional prosody correlation being significant only in V3. In aphasics, emotional prosody correlated also

with chunk length in sung task and in the sung > spoken task difference. Given that emotional prosody perception occurs in a bilateral but right-dominant network (Wildgruber et al., 2006), which is often at least partly preserved in aphasia (Barrett et al., 1999; Ross et al., 1997), this result suggests that the mnemonic benefit of songs on verbal learning in aphasia may be associated also with their emotion-expressing function and the additional emotional cues provided by songs for recall.

In summary, the results of the present study demonstrate that stroke patients benefit from sung repetitive melody as a memory aid in the learning and recall of novel verbal material. However, the cognitive and neural mechanisms underlying the mnemonic effect of songs differ in non-aphasic and aphasic patients, likely owing to differences in lesion patterns and the severity of memory deficits. It seems that in non-aphasic patients, the cues provided by the musical structure facilitate covert rehearsal of the material in verbal working memory, mediated by the left dorsal pathway (AF), resulting in more even recall performance across the verses of the story. Aphasic patients, in turn, seem to benefit from the repetitive melody and rhythm of singing as a means of chunking the words and making the last verse of the story most salient in memory for recall, mediated by bilateral frontal, temporal, and parietal areas as well as the right ventral pathway (IFOF) in particular. One limitation of the present study was the relatively small size of the subgroups (14 aphasic, 17 non-aphasic). In the future, it would be interesting to explore the benefits and mechanisms of musical mnemonics in a larger sample of aphasic patients, including patients with different subtypes and severity levels of aphasia.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2019.101948>.

## Acknowledgements

This work was funded by Academy of Finland (grants 1257077, 277693, 299044), National Doctoral Programme of Psychology, Finnish Cultural Foundation, and Miina Sillanpää Foundation, Finland. We kindly thank the stroke patients for participating in the study and the staff of the Division of Clinical Neurosciences of Turku University Hospital for their help.

## References

- Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—the methods. *Neuroimage* 11, 805–821.
- Atkinson, R.C., Shiffrin, R.M., 1968. Human memory: A proposed system and its control processes. In: Spence, K.W., Spence, J.T. (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory*. 2. Academic Press, New York, pp. 89–195.
- Barnes, J., Ridgway, G.R., Bartlett, J., Henley, S.M., Lehmann, M., Hobbs, N., et al., 2010. Head size, age and gender adjustment in MRI studies: a necessary nuisance? *Neuroimage* 53, 1244–1255.
- Barrett, A.M., Crucian, G.P., Raymer, A.M., Heilman, K.M., 1999. Spared comprehension of emotional prosody in a patient with global aphasia. *Neuropsychiatry*. *Neuropsychol. Behav. Neurol.* 12, 117–1120.
- Belin, P., Van Eeckhout, P., Zilbovicius, M., Remy, P., François, C., Guillaume, S., et al., 1996. Recovery from nonfluent aphasia after melodic intonation therapy: a PET study. *Neurology* 47, 1504–1511.
- Bor, D., Duncan, J., Wiseman, R.J., Owen, A.M., 2003. Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron* 37, 361–367.
- Bor, D., Cumming, N., Scott, C.E., Owen, A.M., 2004. Prefrontal cortical involvement in verbal encoding strategies. *Eur. J. Neurosci.* 19, 3365–3370.
- Bosseler, A.N., Teinonen, T., Tervaniemi, M., Huottilainen, M., 2016. Infant-directed speech enhances statistical learning in newborn infants: an ERP study. *PLoS One* 11, e0162177.
- Breier, J.I., Randle, S., Maher, L.M., Papanicolaou, A.C., 2010. Changes in maps of language activity activation following melodic intonation therapy using magnetoencephalography: two case studies. *J. Clin. Exp. Neuropsychol.* 32, 309–314.
- Brodsky, M.B., McNeil, M.R., Doyle, P.J., Fosset, T.R., Timm, N.H., Park, G.H., 2003. Auditory serial position effects in story retelling for non-brain-injured participants and persons with aphasia. *J. Speech Lang. Hear. Res.* 46, 1124–1137.
- Brown, S., 2017. A joint prosodic origin of language and music. *Front. Psychol.* 8, 1894.
- Buchsbaum, B.R., Olsen, R.K., Koch, P., Berman, K.F., 2005. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron* 48, 687–697.
- Buchsbaum, B.R., Ye, D., D'Esposito, M., 2011. Recency effects in the inferior parietal lobe during verbal recognition memory. *Front. Hum. Neurosci.* 5, 59.
- Callan, D.E., Tsytsev, V., Hanakawa, T., Callan, A.M., Katsuhara, M., Fukuyama, H., et al., 2006. Song and speech: brain regions involved with perception and covert production. *Neuroimage* 31, 1327–1342.
- Calvert, S.L., Tart, M., 1993. Song versus verbal forms for very-long-term, long-term, and short-term verbatim recall. *J. Appl. Dev. Psychol.* 14, 245–260.
- Chiou, K.S., Genova, H.M., Chiaravalloti, N.D., 2016. Structural white matter differences underlying heterogeneous learning abilities after TBI. *Brain Imag. Behav.* 10, 1274–1279.
- Colby, J.B., Soderberg, L., Lebel, C., Dinov, I.D., Thompson, P.M., Sowell, E.R., 2012. Along-tract statistics allow for enhanced tractography analysis. *Neuroimage* 59, 3227–3242.
- Cremers, L.G., de Groot, M., Hofman, A., Krestin, G.P., van der Lugt, A., Niessen, W.J., et al., 2016. Altered tract-specific white matter microstructure is related to poorer cognitive performance: the Rotterdam study. *Neurobiol. Aging* 39, 108–117.
- Crowder, R.G., Greene, R.L., 2000. Serial learning: Cognition and behavior. In: Tulving, E., Craik, F.I. (Eds.), *The Oxford Handbook of Memory*. Oxford University Press, Oxford, England, pp. 125–135.
- Crowder, R.G., Serafine, M.L., Repp, B., 1990. Physical interaction and association by contiguity in memory for the words and melodies of songs. *Mem. Cogn.* 18, 469–476.
- De Renzi, E., Faglioni, P., 1978. Normative data and screening power of a shortened version of the token test. *Cortex* 14, 41–49.
- Deese, J., Kaufman, R.A., 1957. Serial effects in recall of unorganized and sequentially organized verbal material. *J. Exp. Psychol.* 54, 180–187.
- Dick, A.S., Tremblay, P., 2012. Beyond the arcuate fasciculus: consensus and controversy in the connective anatomy of language. *Brain* 135, 3529–3550.
- Duffau, H., Herbet, G., Moritz-Gasser, S., 2013. Toward a pluri-component, multimodal, and dynamic organization of the ventral semantic stream in humans: lessons from stimulation mapping in awake patients. *Front. Syst. Neurosci.* 7, 44.
- Düzel, E., Hufnagel, A., Helmstaedter, C., Elger, C., 1996. Verbal working memory components can be selectively influenced by transcranial magnetic stimulation in patients with left temporal lobe epilepsy. *Neuropsychologia* 34, 775–783.
- Eom, B., Sung, J.E., 2016. The effects of sentence repetition-based working memory treatment on sentence comprehension abilities in individuals with aphasia. *Am. J. Speech Lang. Pathol.* 25, S823–S838.
- Ferreri, L., Verga, L., 2016. Benefits of music on verbal learning and memory: how and when does it work? *Music. Percept.* 34, 167–182.
- Ferreri, L., Bigand, E., Bard, P., Bugaiska, A., 2015. The influence of music on prefrontal cortex during episodic encoding and retrieval of verbal information: a multichannel fNIRS study. *Behav. Neurol.* 2015, 707625.
- Forkel, S.J., Thiebaut de Schotten, M., Dell'Acqua, F., Kalra, L., Murphy, D.G., Williams, S.C., et al., 2014. Anatomical predictors of aphasia recovery: a tractography study of bilateral perisylvian language networks. *Brain* 137, 2027–2039.
- Freebody, P., Anderson, R.C., 1986. Serial position and rated importance in the recall of text. *Discourse Process.* 9, 31–36.
- Frühholz, S., Gschwind, M., Grandjean, D., 2015. Bilateral dorsal and ventral fiber pathways for the processing of affective prosody identified by probabilistic fiber tracking. *Neuroimage* 109, 27–34.
- Gobet, F., Lane, P.C., Croker, S., Cheng, P.C., Jones, G., Oliver, I., et al., 2001. Chunking mechanisms in human learning. *Trends Cogn. Sci.* 5, 236–243.
- Goodglass, H., Kaplan, E., 1983. *The Assessment of Aphasia and Related Disorders*, 2nd ed. Lea & Febiger, Philadelphia.
- Hall, S., Bornstein, R.A., 1991. Serial-position effects in paragraph recall following mild closed-head injury. *Percept. Mot. Skills* 72, 1295–1298.
- Halwani, G.F., Loui, P., Rüber, T., Schlaug, G., 2011. Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Front. Psychol.* 2, 156.
- Haslam, C., Cook, M., 2002. Striking a chord with amnesic patients: evidence that song facilitates memory. *Neurocase* 8, 453–465.
- Hausen, M., Torppa, R., Salmela, V.R., Vainio, M., Särkämö, T., 2013. Music and speech prosody: a common rhythm. *Front. Psychol.* 4, 566.
- Hébert, S., Racette, A., Gagnon, L., Peretz, I., 2003. Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain* 126, 1838–1850.
- Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Innocenti, I., Cappa, S.F., Feurra, M., Giovannelli, F., Santarnecchi, E., Bianco, G., et al., 2013. TMS interference with primacy and recency mechanisms reveals bimodal episodic encoding in the human brain. *J. Cogn. Neurosci.* 25, 109–116.
- Isarida, T., Isarida, T.K., 2006. Influences of environmental context on the recency effect in free recall. *Mem. Cogn.* 34, 787–794.
- Ivanova, M.V., Dragoy, O., Kuptsova, S.V., Yu Akinina, S., Petrushevskii, A.G., Fedina, O.N., et al., 2018. Neural mechanisms of two different verbal working memory tasks: a VLSM study. *Neuropsychologia* 115, 25–41.
- Jefferies, E., Hoffman, P., Jones, R., Ralph, M.A., 2008. The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: a comparative study. *J. Mem. Lang.* 58, 66–87.
- Jungblut, M., Huber, W., Mais, C., Schnitker, R., 2014. Paving the way for speech: voice-training-induced plasticity in chronic aphasia and apraxia of speech—three single cases. *Neural Plasticity* 2014, 841982.
- Kasdan, A., Kiran, S., 2018. Please don't stop the music: Song completion in patients with aphasia. *J. Commun. Disord.* 75, 72–86.
- Kilgour, A.R., Jakobson, L.S., Cuddy, L.L., 2000. Music training and rate of presentation as mediators of text and song recall. *Mem. Cogn.* 28, 700–710.
- Konoike, N., Kotozaki, Y., Miyachi, S., Miyachi, C.M., Yomogida, Y., Akimoto, Y., et al., 2012. Rhythm information represented in the fronto-parieto-cerebellar motor system.

- Neuroimage 63, 328–338.
- Krogsrud, S.K., Fjell, A.M., Tamnes, C.K., Grydeland, H., Due-Tønnessen, P., Bjørnerud, A., et al., 2018. Development of white matter microstructure in relation to verbal and visuospatial working memory – a longitudinal study. *PLoS One* 13, e0195540.
- Laine, M., Goodglass, H., Niemi, J., Koivuselkä-Sallinen, P., Tuomainen, J., Marttila, R., 1993a. Adaptation of the Boston diagnostic aphasia examination and the Boston naming test into Finnish. *Scand. J. Logopedics Phoniatri.* 18, 83–92.
- Laine, M., Tuomainen, J., Ahonen, A., 1993b. Changes in hemispheric brain perfusion elicited by melodic intonation therapy: a preliminary experiment with single photon emission computed tomography (SPECT). *Scand. J. Logopedics Phoniatri.* 19, 19–24.
- Leinonen, L., Hiltunen, T., Linnankoski, L., Laakso, M.J., 1997. Expression or emotional-motivational connotations with a one-word utterance. *J. Acoust. Soc. Am.* 102, 1853–1863.
- Leo, V., Sihvonen, A.J., Linnavalli, T., Tervaniemi, M., Laine, M., Soynila, et al., 2018. Sung melody enhances verbal learning and recall after stroke. *Ann. N. Y. Acad. Sci.* 1423, 296–307.
- Leroy, F., Glasel, H., Dubois, J., Hertz-Pannier, L., Thirion, B., Mangin, J.F., et al., 2011. Early maturation of the linguistic dorsal pathway in human infants. *J. Neurosci.* 31, 1500–1506.
- Lezak, M.D., Howieson, D.B., Bigler, E.D., Tranel, D., 2012. *Neuropsychological Assessment*, 5th ed. Oxford University Press, New York, NY.
- López-Barroso, D., Catani, M., Ripollés, P., Dell’Acqua, F., Rodríguez-Fornells, A., de Diego-Balaguer, R., 2013. Word learning is mediated by the left arcuate fasciculus. In: *Proceedings of the 110. National Academy of Sciences of the United States of America*, pp. 13168–13173.
- Loui, P., 2015. A dual-stream neuroanatomy of singing. *Music. Percept.* 32, 232–241.
- Ludke, K.M., Ferreira, F., Overy, K., 2014. Singing can facilitate foreign language learning. *Mem. Cogn.* 42, 41–52.
- Mabbott, D.J., Rovet, J., Noseworthy, M.D., Smith, M.L., Rockel, C., 2009. The relations between white matter and declarative memory in older children and adolescents. *Brain Res.* 1294, 80–90.
- Mas-Herrero, E., Marco-Pallares, J., Lorenzo-Seva, U., Zatorre, R.J., Rodríguez-Fornells, A., 2013. Individual differences in music reward experiences. *Music. Percept.* 31, 118–138.
- Maylor, E.A., 2002. Serial position effects in semantic memory: reconstructing the order of verses of hymns. *Psychon. Bull. Rev.* 9, 816–820.
- McDonald, C.R., Ahmadi, M.E., Hagler, D.J., Tecoma, E.S., Iragui, V.J., Gharapetian, L., et al., 2008. Diffusion tensor imaging correlates of memory and language impairments in temporal lobe epilepsy. *Neurology* 71, 1869–1876.
- McElhinney, M., Annett, J.M., 1996. Pattern of efficacy of a musical mnemonic on recall of familiar words over several presentations. *Percept. Mot. Skills* 82, 395–400.
- Méndez Orellana, C.P., Van De Sandt-Koenderman, Saliassi, E., Van Der Meulen, I., Klip, S., Van Der Lugt, A., et al., 2014. Insight into the neurophysiological processes of melodically intoned language with functional MRI. *Brain Behav.* 4, 615–625.
- van der Meulen, I., van de Sandt-Koenderman, W.M., Heijnenbrok-Kal, M.H., Visch-Brink, E.G., Ribbers, G.M., 2014. The efficacy and timing of melodic intonation therapy in subacute aphasia. *Neurorehabil. Neural Repair* 28, 536–544.
- Moussard, A., Bigand, E., Belleville, S., Peretz, I., 2014. Learning sung lyrics aids retention in normal aging and Alzheimer’s disease. *Neuropsychol. Rehabil.* 24, 894–917.
- Newhouse, R.C., Holen, M.C., 1975. Individual difference in serial position effects with connected discourse. *Psychol. Rep.* 37, 824.
- Ostergaard, A.L., Meudell, P.R., 1984. Immediate memory span, recognition memory for subspan series of words, and serial position effects in recognition memory for supraspan series of verbal and nonverbal items in Broca’s and Wernicke’s aphasia. *Brain Lang.* 22, 1–13.
- Özdemir, E., Norton, A., Schlaug, G., 2006. Shared and distinct neural correlates of singing and speaking. *Neuroimage* 33, 628–635.
- Palisson, J., Roussel-Baclet, C., Maillet, D., Belin, C., Ankri, J., Narme, P., 2015. Music enhances verbal episodic memory in Alzheimers disease. *J. Clin. Exp. Neuropsychol.* 37, 503–517.
- Peretz, I., Champod, A.S., Hyde, K., 2003. Varieties of musical disorders. In: *The Montreal Battery of Evaluation of Amusia*. 999. *Annals of the New York Academy of Sciences*, pp. 58–75.
- Racette, A., Peretz, I., 2007. Learning lyrics: to sing or not to sing? *Mem. Cogn.* 35, 242–253.
- Racette, A., Bard, C., Peretz, I., 2006. Making non-fluent aphasics speak: sing along!. *Brain* 129, 2571–2784.
- Rainey, D.W., Larsen, J.D., 2002. The effect of familiar melodies on initial learning and long-term memory for unconnected text. *Music. Percept.* 20, 173–186.
- Reggente, N., Cohen, M.S., Zheng, Z.S., Castel, A.D., Knowlton, B.J., Rissman, J., 2018. Memory recall for high reward value items correlates with individual differences in white matter pathways associated with reward processing and fronto-temporal communication. *Front. Hum. Neurosci.* 12, 241.
- Ross, E.D., Thompson, R.D., Yenkosky, J., 1997. Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain Lang.* 56, 27–54.
- Royal, I., Vuvan, D.T., Zengel, B.R., Robitaille, N., Schönwiesner, M., Peretz, I., 2016. Activation in the right inferior parietal lobe reflects the representation of musical structure beyond simple pitch discrimination. *PLoS One* 11, e0155291.
- Rubin, D.C., 1977. Very long-term memory for prose and verse. *J. Verbal Learn. Verbal Behav.* 16, 611–621.
- Salmi, J., Koistinen, O.P., Glerean, E., Jylänki, P., Vehtari, A., Jääskeläinen, I.P., et al., 2017. Distributed neural signatures of natural audiovisual speech and music in the human auditory cortex. *Neuroimage* 157, 108–117.
- Sammler, D., Grosbras, M.H., Anwander, A., Bestelmeyer, P.E., Belin, P., 2015. Dorsal and ventral pathways for prosody. *Curr. Biol.* 25, 3079–3085.
- Samson, S., Zatorre, R.J., 1991. Recognition memory for text and melody for songs after unilateral temporal lobe lesion: evidence for dual encoding. *J. Exp. Psychol. Learn. Mem. Cogn.* 17, 793–804.
- Särkämö, T., Sihvonen, A.J., 2018. Golden oldies and silver brains: deficits, preservation, learning, and rehabilitation effects of music in ageing-related neurological disorders. *Cortex* 109, 104–123.
- Särkämö, T., Tervaniemi, M., Soynila, S., Autti, T., Silvennoinen, H.M., Laine, M., Hietanen, M., 2009. Cognitive deficits associated with acquired amusia after stroke: a neuropsychological follow-up study. *Neuropsychologia* 47, 2642–2651.
- Sarubbo, S., De Benedictis, A., Maldonado, I.L., Basso, G., Duffau, H., 2013. Frontal terminations for the inferior fronto-occipital fascicle: anatomical dissection, DTI study and functional considerations on a multi-component bundle. *Brain Struct. Funct.* 218, 21–37.
- Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., Weiller, C., 2006. Dynamics of language reorganization after stroke. *Brain* 129, 1371–1384.
- Schlaug, G., Marchina, S., Norton, A., 2008. From singing to speaking: why singing may lead to recovery of expressive language function in patients with Broca’s aphasia. *Music. Percept.* 25, 315–323.
- Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., Kolinsky, R., 2008. Songs as an aid for language acquisition. *Cognition* 106, 975–983.
- Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J.L., et al., 2010. Similar cerebral networks in language, music and song perception. *Neuroimage* 51, 450–461.
- Segado, M., Hollinger, A., Thibodeau, J., Penhune, V., Zatorre, R.J., 2018. Partially overlapping brain networks for singing and cello playing. *Front. Neurosci.* 12, 351.
- Sihvonen, A., Ripollés, P., Leo, V., Rodríguez-Fornells, A., Soynila, S., Särkämö, T., 2016. Neural basis of acquired amusia and its recovery after stroke. *J. Neurosci.* 36, 8872–8881.
- Sihvonen, A., Särkämö, T., Ripollés, P., Leo, V., Saunavaara, J., Parkkola, R., Rodríguez-Fornells, A., Soynila, S., 2017a. Functional neural changes associated with acquired amusia across different stages of recovery after stroke. *Sci. Rep.* 7, 11390.
- Sihvonen, A.J., Ripollés, P., Rodríguez-Fornells, A., Soynila, S., Särkämö, T., 2017b. Revisiting the neural basis of acquired amusia: structural changes underlying amusia recovery. *Front. Neurosci.* 11, 426.
- Sihvonen, A.J., Ripollés, P., Särkämö, T., Leo, V., Rodríguez-Fornells, A., Saunavaara, J., Parkkola, R., Soynila, S., 2017c. Tracing the neural basis of music: deficient structural connectivity underlying acquired amusia. *Cortex* 97, 255–273.
- Silverman, M.J., 2007. The effect of paired pitch, rhythm, and speech on working memory as measured by sequential digit recall. *J. Music. Ther.* 44, 415–427.
- Simmons-Stern, N.R., Budson, A.E., Ally, B.A., 2010. Music as a memory enhancer in patients with Alzheimer’s disease. *Neuropsychologia* 48, 3164–3167.
- Spalletta, G., Cravello, L., Gianni, W., Piras, F., Iorio, M., Cacciari, C., et al., 2016. Homotaurine effects on hippocampal volume loss and episodic memory in amnesic mild cognitive impairment. *J. Alzheimers Dis.* 50, 807–816.
- Staffaroni, A.M., Melrose, R.J., Leskin, L.P., Riskin-Jones, H., Harwood, D., Mandelkern, M., et al., 2017. The functional neuroanatomy of verbal memory in Alzheimer’s disease: [18F]-Fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) correlates of recency and recognition memory. *J. Clin. Exp. Neuropsychol.* 39, 682–693.
- Straube, T., Schulz, A., Geipel, K., Mentzel, H.-J., Miltner, W.H.R., 2008. Dissociation between singing and speaking in expressive aphasia: the role of song familiarity. *Neuropsychologia* 46, 1505–1512.
- Su, M., Thiebaut de Schotten, M., Zhao, J., Song, S., Zhou, W., Gong, G., et al., 2018. Vocabulary growth rate from preschool to school-age years is reflected in the connectivity of the arcuate fasciculus in 14-year-old children. *Dev. Sci.* 21, e12647.
- Tamminen, J., Rastle, K., Darby, J., Lucas, R., Williamson, V.J., 2017. The impact of music in learning and consolidation of novel words. *Memory* 25, 107–121.
- Thaut, M.H., Peterson, D.A., McIntosh, G.C., 2005. Temporal entrainment of cognitive functions: musical mnemonics induce brain plasticity and oscillatory synchrony in neural networks underlying memory. *Ann. N. Y. Acad. Sci.* 1060, 243–254.
- Thaut, M.H., Peterson, D.A., McIntosh, G.C., Hoemberg, V., 2014. Music mnemonics aid verbal memory and induce learning-related brain plasticity in multiple sclerosis. *Front. Hum. Neurosci.* 8, 395.
- Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L.W., Dehaene, S., 2014. Learning to read improves the structure of the arcuate fasciculus. *Cereb. Cortex* 24, 989–995.
- Thompson, W.F., Marin, M.M., Stewart, L., 2012. Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19027–19032.
- Wallace, W.T., 1994. Memory for music: effect of melody on recall of text. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1071–1085.
- Wan, C.Y., Zheng, X., Marchina, S., Norton, A., Schlaug, G., 2014. Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca’s aphasia. *Brain Lang.* 136, 1–7.
- Wildgruber, D., Ackermann, H., Kreifelts, B., Ethofer, T., 2006. Cerebral processing of linguistic and emotional prosody: fMRI studies. *Prog. Brain Res.* 156, 249–268.
- Wilson, B.A., Cockburn, J., Baddeley, A., 1985. *The Rivermead Behavioral Memory Test*. Thames Valley Company, Bury St-Edmunds, England.
- Zipse, L., Norton, A., Marchina, S., Schlaug, G., 2012. When right is all that is left: plasticity of right-hemisphere tracts in a young aphasic patient. *Ann. N. Y. Acad. Sci.* 1252, 237–245.
- Zumbansen, A., Peretz, I., Hébert, S., 2014. The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca’s aphasia. *Front. Hum. Neurosci.* 8, 592.