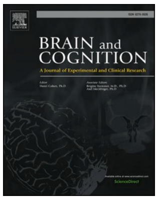
Brain and Cognition 161 (2022) 105881

Contents lists available at ScienceDirect

Brain and Cognition

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



Tonal structures benefit short-term memory for real music: Evidence from non-musicians and individuals with congenital amusia

Yohana Lévêque [a,b,*], Philippe Lalitte [c,d], Lesly Fornoni [a,b], Agathe Pralus [a,b], Philippe Albouy [e,f], Patrick Bouchet [a,b], Anne Caclin [a,b,1], Barbara Tillmann [a,b,1,*]

[a] Lyon Neuroscience Research Center, CNRS, UMR5292, INSERM U1028, Lyon F-69000, France
[b] University Lyon 1, Lyon F-69000, France
[c] CNRS, UMR5022, Laboratoire d'Etude de l'Apprentissage et du Développement, Université de Bourgogne, Dijon, France
[d] CNRS UMR8223, Institut de Recherche en Musicologie (IReMus), Sorbonne Université, France
[e] CERVO Brain Research Center, School of Psychology, Laval University, Quebec, QC, Canada
[f] International Laboratory for Brain, Music and Sound Research (BRAMS) - Centre for Research on Brain, Language and Music (CRBLM), Montreal, QC, Canada

ARTICLE INFO

Keywords:
Tone deafness
Auditory short-term memory
Working memory
Pitch
Implicit processing
Musical structure

ABSTRACT

Congenital amusia is a neurodevelopmental disorder of music processing, which includes impaired pitch memory, associated to abnormalities in the right fronto-temporal network. Previous research has shown that tonal structures (as defined by the Western musical system) improve short-term memory performance for short tone sequences (in comparison to atonal versions) in non-musician listeners, but the tonal structures only benefited response times in amusic individuals. We here tested the potential benefit of tonal structures for short-term memory with more complex musical material. Congenital amusics and their matched non-musician controls were required to indicate whether two excerpts were the same or different. Results confirmed impaired performance of amusic individuals in this short-term memory task. However, most importantly, both groups of participants showed better memory performance for tonal material than for atonal material. These results revealed that even amusics' impaired short-term memory for pitch shows classical characteristics of short-term memory, that is the mnemonic benefit of structure in the to-be-memorized material. The findings show that amusic individuals have acquired some implicit knowledge of regularities of their culture, allowing for implicit processing of tonal structures, which benefits to memory even for complex material.

1. Introduction

Congenital amusia is a lifelong disorder of music perception and production that has been estimated to affect about 1–2% of the general population (Peretz, 2016; Tillmann et al., 2015, for reviews). This deficit cannot be explained by peripheral hearing loss, brain lesions, or general cognitive or social impairments (Ayotte et al., 2002). A first hypothesis attributed this condition to impaired fine-grained pitch processing, with amusic individuals exhibiting elevated pitch discrimination thresholds (Ayotte et al., 2002; Foxton et al., 2004; Hyde & Peretz, 2004). Later studies, however, consistently reported impaired short-term memory for pitch, even in the absence of elevated pitch discrimination thresholds or when the to-be-processed pitch changes exceeded amusics' individual pitch discrimination threshold (Albouy et al., 2013; Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010; Williamson & Stewart, 2010; review in Tillmann et al., 2016). Importantly, the short-term memory deficit is restricted to the musical domain and does not affect verbal material (Albouy, Peretz, et al., 2019), including for classical tests such as forward and backward digit spans (Albouy et al., 2013; Tillmann et al., 2009; Williamson & Stewart, 2010). The deficit extends to short-term memory for other spectral features (timbre features), but not to loudness (Graves et al., 2019; Marin et al., 2012). Electrophysiological and (f-)MRI studies on the cerebral underpinnings of congenital amusia have revealed functional and anatomical abnormalities in the Inferior Frontal Gyrus (IFG) as well as in auditory areas in the temporal lobe, especially in the right hemisphere (Albouy, Mattout, et al., 2013; Hyde et al., 2006, 2007, 2011). Critically, an abnormal connectivity has been reported within this fronto-temporal network (Albouy et al., 2015;

---

* Corresponding authors.
E-mail addresses: yohana.leveque@inserm.fr (Y. Lévêque), barbara.tillmann@cnrs.fr (B. Tillmann).
[1] Equally contributing authors.

https://doi.org/10.1016/j.bandc.2022.105881
Received 22 November 2021; Received in revised form 23 May 2022; Accepted 26 May 2022

Albouy, Mattout, et al., 2013; Hyde et al., 2011; Leveque et al., 2016; Loui et al., 2009), a network already identified as underlying pitch perception and memory in non-amusic individuals (Kumar et al., 2016; Zatorre et al., 1994; see Griffiths, 2001 for a review). Early auditory processing was found impaired in amusic individuals, as reflected by delayed and reduced auditory N100m components while listening to a tone sequence to be encoded in memory (Albouy, Mattout, et al., 2013). Retention and retrieval of pitch in short-term memory were also associated with cerebral differences between amusic and control participants, for instance in oscillatory gamma activity during retention in memory and in the amplitude of the evoked response to a to-be-detected change in a musical sequence (Albouy, Mattout, et al., 2013; see Tillmann, Lévêque, et al., 2016 for a review).

The short-term memory deficit was reported with pitch comparison tasks for single tones and tone sequences. Amusics' short-term memory performance decreased more strongly than performance of control participants in the presence of (a) increased retention delay, (b) interfering material in the retention period, and (c) increased memory load (e.g., Gosselin et al., 2009; Williamson et al., 2010). In addition to these classical features of short-term memory related to forgetting or load, another classical feature of short-term memory is related to the observation that structured material leads to improved memory performance. Long-term memory structures have been shown to robustly influence short-term memory with visual or verbal materials. In the visual domain for instance, material that can be reorganized into higher-level groups (chunks) has been shown to be more easily remembered than unstructured material (Bor et al., 2003, 2004). In the verbal domain, the lexicality, frequency of occurrence of verbal items or the statistical characteristics of phonological information have a significant impact on short-term memory for these items in the general population as well as in children with language impairments (Jones et al., 2010; Jones & Macken, 2018). Preexposition to item sequences with transitional probabilities was shown to influence performance in a subsequent short-term memory task: sequences matching the transitional probabilities were better recalled (Botvinick & Bylsma, 2005; Majerus et al., 2012). This line of research thus showed that experience and implicit structural knowledge are the ground for performance in short-term memory, even for novel sequences. Cowan's model (e.g. Cowan, 1999, 2008) also developed the idea that working-memory is the temporary activation of long-term memory representations. In the auditory domain, previous studies have shown that the inherent structure of music, as based on the regularities of the Western tonal system, improves short-term memory performance for musical sequences in musician and non-musician listeners (in comparison to atonal (unstructured) sequences, e.g., Bharucha & Krumhansl, 1983; Dowling, 1991; Schulze et al., 2012).

Previous neuroimaging studies suggest that musical short-term memory tasks recruit the following network (Albouy, Mattout, et al., 2013; Gaab et al., 2003, p. 2; Griffiths & Green, 1999; Platel et al., 1997; Stevens, 2004; Zatorre et al., 1994): the auditory cortex (posterior superior temporal lobe), frontal regions, more specifically the right IFG and the DorsoLateralPreFrontal Cortex (DLPFC), and the Inferior Parietal Lobule, especially if the information is manipulated (working memory). This network is bilateral but activations are predominant in the right hemisphere. The structure of the information, in particular the tonal organization of music, seems to influence this network during encoding and maintenance in memory. Tonal organization is related to a structure implicitly learned by listeners, since childhood, by mere exposure to the tonal music of their culture. The fMRI study of Schulze, Müller and Koelsch (2011) revealed reduced activation of the right IFG during encoding of tonal compared to atonal sequences in musicians in a working memory task. Other studies, as for example the fMRI study of Cheung et al. (2018) in musicians, corroborated this result, showing the involvement of the right IFG when musical regular structures are violated. The right IFG thus seems to be at the crossroad of tonality (musical structure) processing (see also Tillmann et al., 2003, 2006; Koelsch et al., 2005 in non-musicians) and working memory processing. Schulze et al. (2011) also found a stronger activation of IPL for tonal than atonal music during working memory rehearsal. The bilateral IPL was similarly activated more strongly during encoding of structured versus unstructured visuospatial and auditory-verbal material (Bor et al., 2003, 2004). Finally, a region of the right premotor cortex (BA6, Talairach coordinate: 43, 3, 48) has been reported to be activated more strongly during rehearsal of tonal versus atonal music (Schulze et al., 2011). Activation of this region has also been reported when participants are predicting during the course of auditory sequences (Schubotz, 2007; Schubotz et al., 2003) and premotor activation is associated with the processing of music-syntactic information in a variety of tasks (e.g., auditory oddball paradigms, pitch discrimination tasks, serial prediction tasks; see Janata & Grafton, 2003).

Finally, a region of the right premotor cortex (BA6, Talairach coordinate: 43, 3, 48) has been reported to be activated more strongly during rehearsal of tonal versus atonal music (Schulze et al., 2011). Activation of this region has also been reported when participants are predicting during the course of auditory sequences (Schubotz, 2007; Schubotz et al., 2003) and premotor activation is associated with the processing of music-syntactic information in a variety of tasks (e.g., auditory oddball paradigms, pitch discrimination tasks, serial prediction tasks; see Janata & Grafton, 2003).

Among these principal regions involved in short-term memory, the right IFG has been reported as structurally and functionally altered in congenital amusia. As described above, a dysfunction of the connectivity between the right auditory cortex and the IFG seems to be central in the disorder (Albouy et al., 2015; Albouy, Mattout, et al., 2013; Hyde et al., 2011; Loui et al., 2009; Tillmann, Lévêque, et al., 2016) during encoding, maintenance and retrieval of pitch information in memory. Furthermore, a reduced connectivity between the IFG and the right DLPFC has also been described during maintenance in short-term memory of musical material (Albouy, Peretz, et al., 2019). In contrast, parietal areas do not seem to be affected by congenital amusia and could offer a compensatory support for music memory (Albouy, Mattout, et al., 2013).

Tonal structure could provide a support for this impaired musical processing observed in congenital amusia, with deficits in the fronto-temporal pathway crucial for musical working memory (Cheung et al., 2018; Kumar et al., 2016; Schulze et al., 2011; Schulze & Koelsch, 2012). Albouy, Schulze, et al. (2013) indeed provided first evidence for some benefit of tonal structure in the to-be-remembered tone sequences in congenital amusia. Using an artificial material with simple tonal structures (i.e., 5-tone sequences), this benefit was observed only for response times in amusic individuals. No benefit of tonal structure was observed for memory performance (as measured by *d'*) in the amusic group. Only control participants showed this advantage. To our knowledge, no other study has investigated yet whether tonal structures in the material, calling for the tonal knowledge of the listener stored in long-term memory, might support short-term memory performance in congenital amusia. We thus set out in the present study to use more complex musical material, which starts from real tonal, ecologically valid material to bring complementary evidence of the benefit of tonal structure for short-term memory in congenital amusia. Research on musical memory in congenital amusia has essentially used highly-controlled, simple short melodies, created for the study or not, composed of a single melodic line (e.g., Graves et al., 2019; Omigie et al., 2013; Quiroga-Martinez et al., 2021; Weiss & Peretz, 2019). Understanding of musical memory in congenital amusia is thus suffering from lack of data on memory for real-world music. To our knowledge, only one study has used harmonized or orchestrated music to explore musical memory in congenital amusia, but here testing potential knowledge of tonal exemplars stored in long-term memory (Tillmann et al., 2014, investigating familiar music recognition in very short excerpts).

The benefit of tonality in amusia reported by Albouy et al. (2013) integrates in findings of other studies suggesting that also amusic individuals have some knowledge about tonal structures. The hypothesis is that this knowledge is acquired by mere exposure to music that obeys to the rules of the Western tonal musical system, as reported for nonmusician listeners (Bigand & Poulin-Charronnat, 2006). For amusic individuals, data acquired with implicit investigation methods have revealed some knowledge about the syntactic-like functions of chords in the Western musical system (Tillmann et al., 2012; see also Omigie et al., 2012) as well as related to mode (major, minor; Gosselin et al., 2015). Amusic individuals seem thus to be able to acquire tonal knowledge via mere exposure in everyday life thanks to the cognitive capacity of implicit learning, as previously discussed for nonmusician listeners (Bigand & Poulin-Charronnat, 2006; Tillmann et al., 2000). In a related vein, Omigie and Stewart (2011) provided evidence in the laboratory that individuals with amusia can learn tonal regularities of new artificial material via statistical learning (but see Peretz et al., 2012). Recent electroencephalographic data by Quiroga-Martinez et al. (2021) suggested that amusic individuals, just like control participants, show reduced responses to pitch deviants in melodies that are less predictable according to long-term representations (unfamiliar or highly complex melodies) compared to more predictable melodies. Amusic individuals also exhibited processing differences between tonal and atonal music when subjectively evaluating musical material with three questions: an explicit structural one, a personal, emotional one, and a more social one (judging the perception of others). The question type influenced the extent of the structure processing that the measurement can reveal: while amusic individuals were impaired for the question requiring explicit structural judgments, they performed as well as their matched controls for the two other questions, suggesting that amusics possess more tonal knowledge than revealed by tasks requiring explicit judgments (Tillmann, Lalitte, et al., 2016). In this previous study, we created an atonal material as a counterpart of the tonal material; notably by manipulating the pitch content (i.e., removing the characteristic patterns of the tones' frequencies of occurrence) while keeping the same rhythm, tempo, dynamics and thematic structures as in the original tonal versions (procedure adapted from Lalitte et al., 2009). The tonal and atonal versions thus differed only by the used tone sets, resulting in the presence versus absence of tonality.

Based on this tonal and atonal material of Tillmann et al. (2016), we here aimed to test whether it is possible to observe short-term memory improvement for complex tonal material (in comparison to atonal material) even in congenital amusia and whether this tonal structure benefit can also extend to the performance level. In contrast to Albouy et al. (2013) who used only 5-tone sequences (as did Schulze et al., 2012), we used the more complex (real musical) material of Tillmann et al. (2016), which was richer in musical structures and for which we have previously shown that both control and amusic participants can differ between tonal and atonal versions in subjective evaluations. For control participants, we expected to observe a short-term memory advantage for the tonal material (in comparison to the atonal material), as previously reported with more simple tone or chord sequences (Albouy, Schulze, et al., 2013; Bharucha & Krumhansl, 1983; Dowling, 1991; Schulze et al., 2012). As the amusic participants have been shown to be able to perceive differences between these tonal and atonal materials (even though in a longer duration, Tillmann, Lalitte, et al., 2016), we made the hypothesis that they could also benefit from the tonal structure in a short-term memory task, even though their performance would be overall impaired in comparison to the controls. If we observe the tonal structure benefit for short-term memory performance, this finding would provide new evidence that amusic individuals have acquired some long-term knowledge of the tonal system. This knowledge could then also serve for processing and memorizing new input, in particular when the tonal material has a rich musical structure, providing multiple cues related to tonality that to facilitate encoding and retention. More generally, this tonality effect would be new evidence that long-term representations influence short-term memory in the auditory domain, even for novel materials, as it has been robustly shown in the verbal and visual domains. As the material used here was more complex and longer than in previous short-term memory studies using simple tone sequences (e.g., five tones in Albouy et al., 2013), we slightly adapted the delayed matching-to-sample paradigm where participants had to determine whether two musical excerpts were the same or different. Instead of presenting the excerpts in pairs (Sequence1, S1, followed by Sequence 2, S2), the first excerpt was presented twice to facilitate encoding, and was then followed by S2, which could be same or different (i.e., S1 – S1 – S2).

2. Material and methods

2.1. Participants

Twenty-one amusic adults (fourteen women) and 21 matched nonmusician controls (fifteen women) participated in the study (see Table 1 for characteristics of each participant group). The groups did not differ regarding age (p =.31), education level (p =.55) and musical training (p =.23), as tested with two-sided t-tests. In the present study, intellectual efficiency was not evaluated as previous studies have shown that individuals with congenital amusia do not differ from control participants in verbal and non-verbal abilities, except for musical material (Albouy, Mattout, et al., 2013; Ayotte et al., 2002; Foxton et al., 2004; Peretz et al., 2002; Williamson & Stewart, 2010). Furthermore, verbal short-term memory evaluated with a delayed matching to sample task has been shown to be intact in amusics (Albouy, Peretz, et al., 2019; Tillmann et al., 2009), indicating that the task structure does not represent a difficulty for individuals with amusia, and thus cannot explain their impairment when applied to non-musical material. Moderate or severe peripheral hearing loss was excluded using standard audiometry, and all participants reported no history of neurological or psychiatric disease. Participants provided written informed consent and received a monetary compensation for their participation. The study procedures were approved by the appropriate ethics committee (Comité de Protection des Personnes, CPP).

All participants were tested with the MBEA (Peretz et al., 2003). To be considered as amusic, participants had to obtain an average score two standard deviations below the average of the normal population on the MBEA. Amusics obtained scores below the cut-off score on the overall MBEA battery (23.4 on average across the six tasks, maximum score = 30) or the three pitch-related subtests (scale, interval, contour, cut-off = 21.7 on average across the three tasks, maximum score = 30; Liu et al., 2010), except one amusic participant with borderline scores on both measures (23.5 and 22.33). All controls obtained scores higher than the cut-off for both scores. The average scores of the amusic group differed significantly from the scores of the control group for the allover, global MBEA battery score (see Table 1; p <.0001) and the MBEA pitch score (see Table 1; p <.0001). Pitch discrimination thresholds (PDTs) were determined using a two-alternative forced-choice task with an adaptive tracking, two-down/one-up staircase procedure (see Tillmann et al., 2009, for task and details). PDTs of the amusic group were significantly higher than those of the control group (see Table 1; p =.025).

Table 1
Characteristics of amusic and nonmusician control groups.

	Amusics		Controls		
	mean	SD	mean	SD	*t*-test
Age (years)	41.76	16.47	36.98	13.65	p =.31
Education (years)	14.62	2.84	15.10	2.28	p =.55
Musical training (years)	0.29	0.90	0.71	1.35	p =.23
MBEA global score	22.14	1.33	26.81	1.47	p <.0001
MBEA pitch score	20.81	1.70	26.56	1.82	p <.0001
Pitch discrimination threshold (semitone)	0.98	1.08	0.35	0.20	p =.025

2.2. Material

The material was based on the tonal and atonal material of Tillmann et al. (2016): the tonal material was selected from romantic and early twentieth century piano repertory, and the atonal counterpart was based on an entire reorganization of the pitch content (i.e., the used pitches and intervals) of the tonal material with a pseudorandom process with constraints (see Lalitte et al., 2009, for details). Note that the atonal versions shared the same number of notes, rhythm, articulations, tempo, dynamics, and thematic structure with tonal versions. For the present study, we selected 12 short musically meaningful excerpts in the tonal set and their corresponding 12 excerpts in the atonal set (average duration of 6.82 sec ± 1.01; min = 4.80; max = 8.10). See Fig. 1 for an example and Supplementary Online Material for the associated sound files.

All tonal excerpts had a clearly established tonality, while the atonal versions were missing a tonal center. Following Tillmann et al. (2016), we analyzed the sound files with the MIR toolbox (Lartillot & Toiviainen, 2007), which extracts multiple acoustic and musical features from audio files. We retained the same MIR parameters as in Tillmann et al. (2016), notably to cover low-level acoustic-feature information and higher-level, more global, cognitive features related to tonality. For the low-level acoustic-feature information, the following parameters were retained: (1) standard deviation of intensity (as measured by the SD of the root-mean-square energy RMS), (2) mean roughness (or sensory dissonance; based on (Plomp & Levelt, 1965), related to the beating phenomenon when overtones are close in frequency), as amusics have been reported to be sensitive to this acoustic feature (Cousineau et al., 2012; Marin et al., 2015); (3) spectral novelty (based on the similarity of the harmonic spectrum between time point *t* and *t-1*) and spectral flux (related to the rate of change of the spectral shape (Lerch, 2012)), as both are tapping into the pitch dimension, which is manipulated between the tonal and atonal versions. For the higher-level, more global, cognitive features related to tonality, the following parameters were used: (1) key clarity (associated to the average strength of the best fitting keys over time; based on Krumhansl & Jusczyk, 1990), (2) mean Harmonic Change Detection Function (HCDF; flux of the tonal centroid; Harte et al., 2006), (3) chromagram novelty (based on the similarity of the chromagram between time point *t* and *t-1;* the chromagram is defined as the pitch-class distribution based on the number of occurrences of the specific pitch class in a time frame and its energy throughout the analysis block). As in Tillmann et al. (2016), the tonal and atonal versions did not differ significantly on intensity ($p = .13$) and roughness ($p = .67$) and differed only marginally significantly on spectral novelty ($p = .07$). They differed significantly on spectral flux ($p < .001$), key clarity ($p < .0001$), HCDF ($p = .004$), and chromagram novelty ($p = .03$). These acoustic analyses thus confirmed that the short excerpts had overall the same characteristics as those implemented in the longer material used in Tillmann et al. (2016), where the excerpts had an average duration of 26 sec (*SD* = 3.86). In particular, they confirmed that atonal and tonal stimuli were differing by features linked to tonality and not low-level acoustic characteristics.

To construct the “different” trials in the memory task, five to eight notes of the excerpts were modified. This consisted in either a change of register (same notes one octave up or down, see Fig. 1 for an example), a change of melodic contour, or a change of register and melodic contour. Four additional excerpts (two tonal and two atonal ones) were constructed as example trials, with two same and two different trials.

As repetition of information might influence memory performance (i.e., leading to less varying material, thus fewer different information needs to be memorized and/or the same memory trace might be reinforced), we added a further acoustic analysis of our material using the MIR Toolbox. We first calculated the Mel Frequency Cepstrat Coefficients (MFCCs; Lartillot & Toiviainen, 2007) and on its basis, the Novelty curve (Foote, 2000). The novelty curve measures the strongest changes in an audio signal, and is indirectly a measure of repetition (i.e., numerous changes represent less repetition). This measure did not differ between the tonal excerpts (both original and modified ones) and their atonal counterpart excerpts, $p = .40$.

All materials were recorded in MIDI format with Steinberg Cubase SX 2 Software and exported in audio format (aiff) with the Halion Acoustic Grand Piano plug-in, which provides a realistic piano timbre. A small amount of artificial reverberation has been added to give a concert hall effect (Roomwoks SE, Large Living Room). Presentation software (Neurobehavioral Systems, Albany, CA, USA) was used to present the stimuli and to record participants’ responses.

Material validation tests. We ran two short evaluation tests on the entire material set. In the first test, 20 participants (not included in the main experiment) evaluated all excerpts using question 1 of Tillmann et al. (2016), requiring participants to evaluate on a scale from 1 (little) to 10 (strong) the degree the excerpt is in agreement with what we are used to hear as music respecting the musical system of our culture. The ratings confirmed that tonal pieces reached significantly higher ratings (6.92 ± 1.34) than did atonal excerpts (4.47 ± 1.06), $p < .0001$. In addition, we ran item-based analyses, which confirmed this significant difference over all pairs of tonal/atonal excerpts, $p < .0001$. In a second test, 19 participants (with 18 having also performed test 1), listened to the original and modified versions in direct comparison (for both tonal and atonal materials). One of the two versions (original or modified) was presented twice followed by the other version in comparison (i.e., original-original-modified, or modified-modified-original). Participants were informed that the excerpts differed only slightly and were asked to judge the salience of the change between the two melodies on a scale from 1 (weak) to 10 (very strong). Results revealed that changes were perceived as equally strong for tonal (5.18 ± 2.11) and atonal (5.43 ± 2.10) versions ($p = .22$); this was confirmed by an item-based analysis ($p = .23$). In sum, these material validation tests confirmed that tonal and atonal versions were perceived as more or less typical of Western tonal music, as intended (and in agreement with Tillmann et al., 2016), and that the changes between original and modified versions were perceived as equally salient in tonal and atonal items.

2.3. Procedure

Participants were informed that they would listen to musical excerpts. A given excerpt (S1) was presented twice (separated by a silent delay of 500 ms) to allow memorizing it, and then after a silent delay of 2 sec, a second excerpt (S2) was presented. Participants were asked to press one of two mouse buttons to indicate whether this second excerpt was the same as the first excerpt or different from the first excerpt. To clarify the trial presentation, the information “melody 1” was displayed on the screen during the presentations of the first excerpt (S1), and “melody 2” was presented during the presentation of the second excerpt (S2); nothing was presented on the screen during the silent delays. Participants could respond while the second excerpt was playing and after the second excerpt. The next item was started by pressing any mouse button. At the beginning of the experiment, the task was explained with two same trials (one tonal, one atonal) and two different trials (one tonal, one atonal). Error feedback was given only for these practice trials.

The experiment consisted of 48 trials, each excerpt (12 tonal, 12 atonal) was used as S1 once in a same trial and once in a different trial. A pseudo-randomized presentation was used so that: (1) the same excerpt was not presented consecutively in a “same” and a “different” trial, and (2) the type of trial (same/different) changed after at most 3 trials (i.e., no more than three consecutive “same” or “different” trials). For the presentation of the excerpts within the experiment, six different orders were created. In addition, two versions were created for each order: the first excerpts were either the original excerpts or the modified excerpts (i.e., modified for the “different trials”); controlling for potential differences between the excerpts across participants. One of the potential 12 programs was attributed to a pair of participants of each group

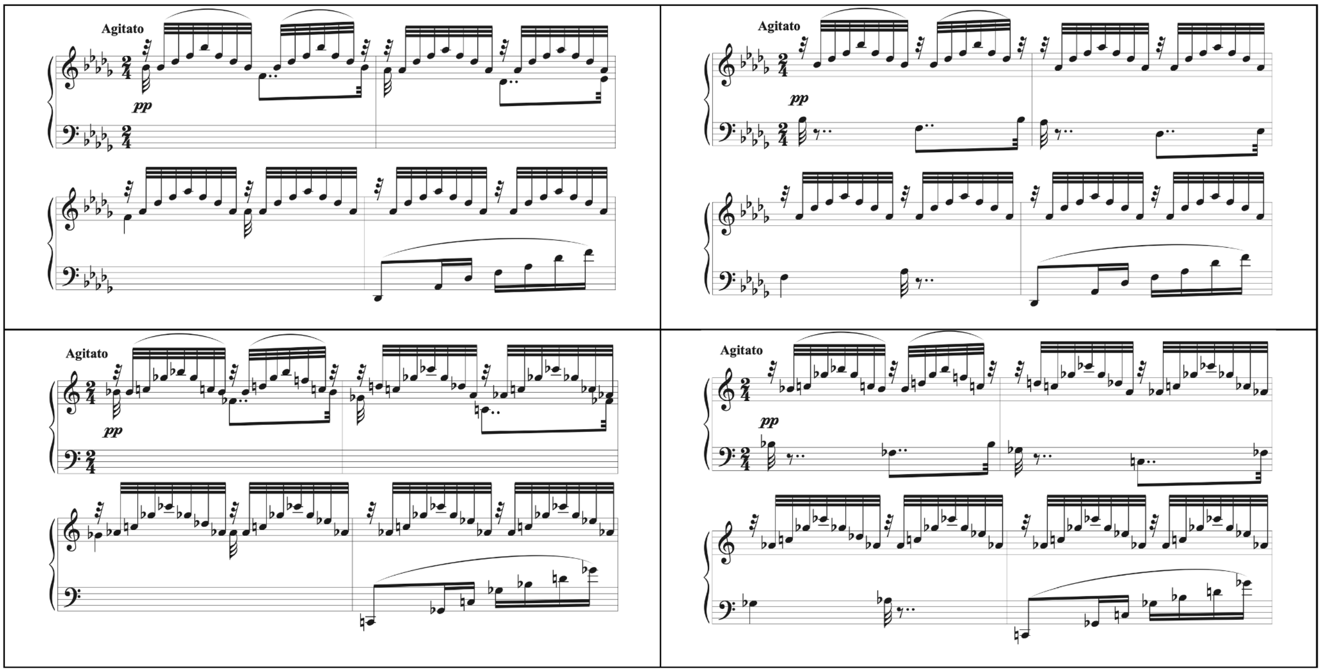


Fig. 1. Scores of one of the tonal excerpts (top left) and its matching atonal counterpart (bottom left) to illustrate the matching of the tonal and atonal versions on various features, such as rhythm, contour, tempo, dynamic and thematic structure. The tonal excerpt is taken from Sinding, C.: Rustle of Spring opus 32n° 3. The right side of the figure represents the different trial associated to the tonal excerpt (top right) and the atonal excerpt (bottom right). The modification was a change of register of eight notes of the left-hand voice, notably one octave down (bars 1 to 3). See supplementary online material for soundfiles of the excerpts.

(amusic, control) in a yoked way, allowing for further control of potential influences of material features.

2.4. Data analyses

Performance was analyzed using signal detection theory by calculating, for each participant and for each condition (tonal or atonal), the discrimination sensitivity (*d'*) and the response bias (*c*)[2]. For each participant, these analyses were based on hit rate (i.e., number of correct responses for different trials / number of different trials) and false alarm rate (i.e., number of incorrect responses for same trials / number of same trials). Positive values for *c* arise when the miss rate (incorrect responses for different trials / number of different trials) exceeds the false alarm rate. Positive values thus indicate a tendency to answer "same", negative values indicate a tendency to answer "different", and *c*-values around 0 suggest the absence of a response bias. *d'* and *c* were analyzed respectively with a 2x2 ANOVA with Group (amusic, control) as the between-participants factor and Tonality (tonal, atonal) as the within-participant factor. To test whether *d'* or *c* was significantly above 0, we used a two-sided *t*-test.

Response Times (RTs) of correct responses were measured from the beginning of S2, and thus reflected decision times for the memory task. For each participant, median RTs were calculated for each condition, for correct trials only. One amusic participant was excluded from RT analysis because in two conditions there were no correct responses (i.e., thus no correct RTs). Correct RTs were analyzed with a 2x2x2 ANOVA with Group (amusics, controls) as the between-participants factor, and Tonality (tonal, atonal) and Trial type (same, different) as the within-participant factors.

A post-hoc statistical power analysis was performed on the analyses using G*Power (Faul et al., 2007, version 3.1.9.4), with Power defined as 1- β error probability. The following input parameters were used: effect size *f* computed based on the Partial η^2, α error probability of 0.05, total sample size of 42 (33 for analyses without data from participants with very low performance), number of groups 2, number of measurements 4 and the correlation among repeated measures. Power computation for correlation analyses was done for one-tailed tests. Power is a value comprised between 0 and 1 and power below 0.80 is usually considered insufficient.

3. Results

3.1. Sensitivity d' and response bias c

For *d'* (Fig. 2), the main effect of Group was significant, $F(1, 40) = 20.012$, $p < .0001$, $MSE = 1.155$, *partial* $\eta^2 = 0.333$, statistical power > 0.99, with control participants reaching higher performance levels than did amusic participants. The main effect of Tonality was also significant, $F(1, 40) = 4.082$, $p = .050$, MSE = 0.264, *partial* $\eta^2 = 0.093$, statistical

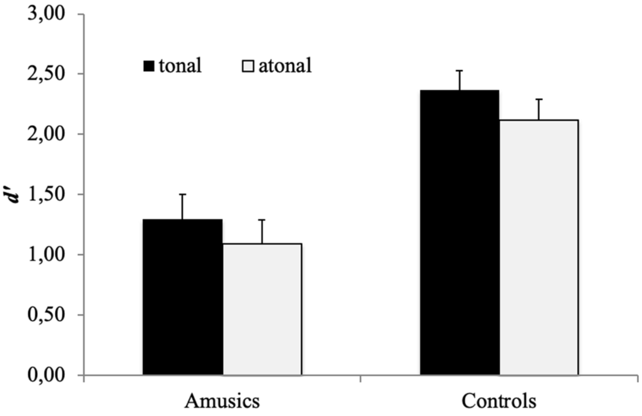


Fig. 2. Average performance *d'* presented as a function of participant Group (amusics, controls) and Tonality (tonal, atonal). Error bars represent between-participant standard errors.

---

[2] 2 The correction of *d'* and *c* measures used 0.01 for cases without false alarms and 0.99 for the maximum number of hits.

power > 0.99 with higher performance levels for tonal excerpts than for atonal excerpts. The interaction between Group and Tonality was not significant, p =.825, *partial* $\eta^2 = 0.001$, with insufficient statistical power (0.09). Performance was above 0 for both groups in each condition, all p-values < 0.0001.

For response bias *c* (Table 2), the main effect of Group was significant, $F(1, 40) = 6.039$, p =.018, $MSE = 0.396$, *partial* $\eta^2 = 0.13$, statistical power = 0.84, with amusic participants showing a positive bias (i.e., a tendency to respond "same" for tonal and atonal sequences), which was significantly different from 0 in each condition (*ps* < 0.05), but not control participants (*c* did not differ from 0 in each condition (*ps* > 0.321). The main effect of Tonality failed to reach significance, $F(1, 40) = 3.72$, p =.061, MSE = 0.117, partial $\eta^2 = 0.085$, statistical power = 0.64, and the interaction was not significant, p =.830.

3.2. Correct response times

For correct RTs (Fig. 3), the main effect of Group was not significant (p =.098, power = 0.15) but the main effect of Trial type was significant, $F(1, 39) = 27.296$, p =.018, $MSE = 1.823 \times 10^6$, *partial* $\eta^2 = 0.412$, statistical power > 0.99, with faster RTs for different trials. Most importantly, the interaction between Group and Trial type was significant, $F(1, 39) = 9.959$, p =.003, $MSE = 1.823 \times 10^6$, *partial* $\eta^2 = 0.203$, statistical power = 0.88. While RTs did not differ between the two groups for same trials (p =.865), amusic participants responded significantly slower than control participants for different trials ($F(1, 39) = 6.147$, p =.018). There were no other significant main effect or interactions.

3.3. Correlations with MBEA and PDT

To investigate whether musical skills in general could affect performance to our task, beyond the amusia condition, which is a binary variable, we calculated correlations between participants' scores in the MBEA as well as their PDTs with (a) average *d'* (across the tonal and atonal conditions) and (b) the difference of *d'* scores in tonal and atonal conditions (i.e., reflecting the tonality effect). We calculated these correlations 1) across all participants, and 2) for each participant group separately (this is particularly important for correlations involving the MBEA as the MBEA score had served to create the groups, thus correlations across all participants might reflect group differences only). Average *d'* was positively correlated with MBEA score across all participants ($r(40) = 0.757$, p <.001, Power > 0.99, reflecting in part the group effect reported in short-term memory performance), as well as for amusic and control groups considered separately ($r(19) = 0.682$, p <.001, Power = 0.99 and $r(19) = 0.579$, p =.006, Power = 0.93, respectively;). Average *d'* was negatively correlated with PDTs across all participants ($r(40) = -0.360$, p =.019, Power = 0.79, reflecting the group effects on pitch discrimination threshold reported in section 2.1 and on short-term memory performance), but not for amusic and control groups separately ($r(19) = -0.256$, p =.264, and $r(19) = -0.249$, p =.276, respectively; Power = 0.31 and 0.30 respectively). In contrast, all correlations of MBEA or PDT with *d'* differences (tonality effect) were not significant, all *p*-values > 0.163, power below 0.37.

Table 2
Mean and standard errors (SE) for response bias *c*, presented as a function of tonal and atonal conditions for amusic and control participants.

	tonal		atonal	
	mean	SE	mean	SE
amusics	0.427	−0.115	0.299	0.130
controls	0.106	0.105	−0.055	0.088

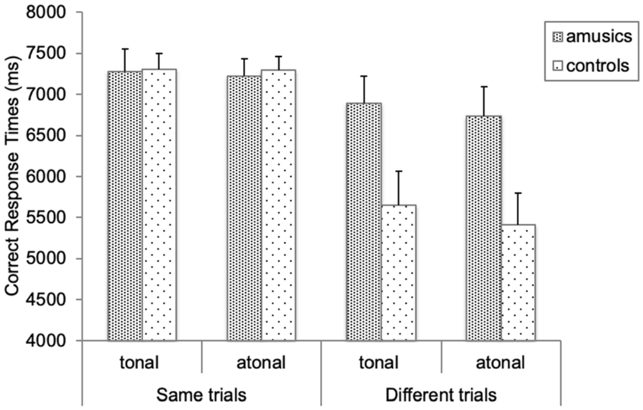


Fig. 3. Correct response times (in ms) presented as a function of participant Group (amusics, controls), Trial type (same, different), and Tonality (tonal, atonal). Error bars represent between-participant standard errors.

3.4. Excluding participants with low performance

As performance for some of the amusics was rather low, we screened performance of all participants for average *d'* performance inferior to 1. We thus excluded eight amusic participants and one control participant and analyzed *d'*, *c*, and RT data with the reduced group (13 amusic and 20 control participants, Table 3). For *d'*, this reduced group analyses confirmed the main effects of Group ($F(1, 31) = 8.425$, p =.007, $MSE = 0.968$, *partial* $\eta^2 = 0.214$, power = 0.92) and of Tonality ($F(1, 31) = 5.788$, p =.022, $MSE = 0.266$, *partial* $\eta^2 = 0.157$, power > 0.99), and no interaction (p =.950, *partial* $\eta^2 < 0.001$, power = 0.05). For *c*, this reduced group analysis confirmed the main effect of group, but it fell short of significance, $F(1, 31) = 3.319$, p =.078, $MSE = 0.230$, *partial* $\eta^2 = 0.097$, power = 0.97. The main effect of Tonality and its interaction with Group were not significant (p =.130 and p =.907, respectively). For correct RTs, the analysis confirmed the significant main effect of Trial type, $F(1, 31) = 23.096$, p <.0001, $MSE = 2.119 \times 10^6$, *partial* $\eta^2 = 0.427$, power > 0.99, and its interaction with Group, $F(1, 31) = 5.489$, p =.026, $MSE = 2.119 \times 10^6$, *partial* $\eta^2 = 0.151$, power > 0.99.

4. Discussion

Our study investigated whether tonal structure of to-be-memorized tone material might benefit short-term memory performance in amusic participants and non-musician control participants with real-world music material. By comparing short-term memory of tonally structured and unstructured musical material, we aimed to provide further

Table 3
Mean and standard errors (SE) for *d'*, response bias *c* and correct response times (RT) in ms for the reduced analysis including only participants performing superior to a *d'* of 1 (13 amusic participants, 20 control participants), presented as a function of tonal and atonal conditions for amusic and control participants.

		tonal		atonal	
		mean	SE	mean	SE
d'	amusics	1.752	0.230	1.432	0.180
	controls	2.464	0.184	2.159	0.169
bias c	amusics	0.307	0.102	0.152	0.120
	controls	0.076	0.103	−0.058	0.090
RT (same)	amusics	7312.308	313.092	7050.385	282.443
	controls	7303.100	203.272	7282.100	177.331
RT (different)	amusics	6647.308	445.570	6438.154	476.275
	controls	5539.200	413.181	5338.950	400.709

evidence for amusics' implicit tonal knowledge and its impact on short-term memory, in particular on performance level (and not only on response times, as shown in Albouy, Schulze, et al., 2013).

Our data confirmed amusic individuals' short-term memory deficit in comparison to control participants, here documented for the first time using realistic musical material. Most importantly, memory performance was better for tonal material than for atonal material in *both* participant groups. In agreement with previous data on tone and chord material (Bharucha & Krumhansl, 1983; Dowling, 1991; Schulze et al., 2012), tonal structure led to improved short-term memory performance in non-musician listeners, here observed with more complex musical material. Furthermore, the results obtained here extend the previous findings of non-amusic participants (including the control participants of Albouy et al., 2013) to amusic individuals. In contrast to Albouy et al. (2013) who used simple sequences of five tones presented isochronously, we here used real musical excerpts as the tonal versions and created similarly complex, highly comparable atonal versions thereof. The use of complex musical excerpts together with an increased duration (i.e., 7 s on average in contrast to 2.5 s in Albouy et al., 2013) and increased tone information made the task more difficult, leading us to adapt the delayed-matched to sampling paradigm to a new presentation format (i.e., S1 – S1 – S2). Note that the task difficulty is reflected in the performance level of the controls: for the tonal material in Albouy et al. (2013), control participants reached an average *d'* of 3.45 while they only reached an average *d'* of 2.37 with the present material. This increased overall difficulty of the task did not prevent to observe tonality effects, in both participant groups, emphasizing the robustness of the tonality effect with a musically rich material.

In addition to enhanced memory load, using real-world orchestrated music, compared to a poorer musical material, may have: (i) increased pleasure and motivation, thus attention, (ii) complexified the harmonic organization (higher-order relationships between notes), which means that participants had more complex rules to mobilize to predict and optimize the processing of the excerpts. This material better reflects encoding of musical pieces amusic individuals are exposed to in daily life and the difficulties they report in memorizing tunes from their social environment repertoire. For instance, in the pioneer study by Ayotte, Peretz & Hyde (2002), amusic participants were impaired in identifying familiar folk songs when listening to the melodies without lyrics, while they could name the title with an excerpt of the lyrics. Note that one major difference between our experimental context and musical exposition in daily life is that encoding is principally implicit in daily life, and was voluntary, in response to instructions, in our design. Interestingly, in Ayotte et al.'s study (2002), detection of tone deviants was better in familiar than in unfamiliar melodies, revealing that amusic individuals have long-term memory knowledge of musical exemplars (see also Graves et al., 2019; Quiroga-Martinez et al., 2021; Tillmann et al., 2014, in congenital amusia; and Besson & Faïta, 1995, for controls). In the present study, we show that participants have long-term memory knowledge on structures of the tonal system (see also Tillmann et al., 2012; Tillmann, Lalitte, et al., 2016) and that this type of long-term memory knowledge can also influence short-term memory performance in the amusic population as it does in control populations.

Despite the repeated presentation of the S1 stimulus, amusic individuals' performance was below controls' performance, confirming amusics' overall short-term memory deficit for musical material. Thus, even though amusics showed a benefit for tonal material over atonal material, this tonal benefit did not allow to fully overcome their short-term memory limits for pitch material. While amusic participants' overall performance level (as measured by the average of d' across tonal and atonal conditions) correlated with their MBEA scores, there was no correlation between MBEA scores and the tonality effect (the size of the difference in performance between tonal and atonal conditions), and this tonality effect was not different from that of control participants. The correlations computed should be interpreted with caution given our group sample size ($n = 21$)[3]. Nonetheless, this observation points to some degree of dissociation between performance in explicit tasks with musical material (such as the short-term memory task used here or in some sub-tests of the MBEA), which is impaired in congenital amusia, and facilitatory effects of musical (tonal) structure, which are preserved in congenital amusia (e.g., Tillmann, Lalitte, et al., 2016). This pattern of results is also similar to findings in the linguistic domain for another pathological population. For instance, children with developmental language impairment showed reduced verbal short-term memory/working memory performance in an oral repetition task, but their performance was strongly influenced by language-related long-term memory effects (Jones et al., 2010). Our study brings new evidence that long-term memory structures influence performance in the musical domain as it has been robustly shown in the verbal and visual domains (e.g., Bor et al., 2003; Jones & Macken, 2018; Tulving & Patkau, 1962).

Thus, congenital amusics' short-term memory shows not only classical features related to load and forgetting (e.g., Gosselin et al., 2009; Williamson et al., 2010; Williamson & Stewart, 2010), but also related to benefits of structure in the to-be-memorized material. This is in line with the observation that amusic individuals are able to learn tonal rules using statistical learning (Omigie & Stewart, 2011) and have some implicit tonal knowledge, as shown by intact harmonic priming effect (Tillmann et al., 2012), subjective judgements favoring tonal over atonal excerpts (Tillmann, Lalitte, et al., 2016) or shorter response times for tonal than for atonal melodies (Albouy, Schulze, et al., 2013). This structural benefit was also previously observed for non-amusic listeners (both musicians and nonmusicians) for tonally structured music or music-like materials (Albouy, Schulze, et al., 2013; Bharucha & Krumhansl, 1983; Dowling, 1991; Schulze et al., 2012). Thus, tonal knowledge in long-term memory (referred to as schematic knowledge by Bharucha, 1987) robustly influences the way we predict the course of a musical sentence. The study by Justus and Bharucha (2001) showed, for instance, that response time differences in a harmonic priming paradigm still reflected the expectations based on representations in long-term memory of tonal relationships even when new information was repeatedly presented, and thus present in short-term memory (see also Filipic et al., 2010; Guo & Koelsch, 2016; Koelsch & Jentschke, 2008).

Using real-world complex material also calls for controls in material across conditions. The feature analyses we performed with MIR on our stimuli allowed us to discard an explanation in terms of low-level acoustic differences between the tonal and atonal conditions (intensity, roughness). Atonal material was not just a more rough or dissonant material, leading to worse short-term memory performance. Furthermore, an alternative interpretation linked to differences in musical repetition within excerpts could also be discarded by our analyses. Our feature analyses confirmed that differences were linked to tonality: spectral flux, key clarity and chromagram novelty distinguished the two conditions. Listeners' long-term memory knowledge about tonality might help for structuring and memorizing incoming information. In tonal material, the use of tones respects frequencies of occurrence and co-occurrence as formalized by the Western tonal system. Exploiting these regularities with knowledge about possible structures (including harmonic structures or implicit harmony), which are stored in long-term memory, might help decreasing the amount of to-be-stored information (e.g., via chunking; Gobet et al., 2001; Miller,

---

[3] 3 Note however that samples in previous studies were frequently smaller than in the present study (e.g., n=15 in Omigie et al., 2013; n=13 in Gosselin et al., 2015; n=11 in Tillmann, Lalitte, et al., 2016) and that the power reached here was >0.93 for these correlations between MBEA and *d'* in each subgroup.

1956) or providing a grid for detecting changes when listening to S2. Further research is needed to determine the exact mechanisms underlying the benefit of structure in tonal material. It would also be interesting to test how much familiarity with musical exemplars could also be a factor influencing short-term memory performance in amusia, even when explicit recognition is deficient. As some studies with implicit methods of evaluation have shown that amusic individuals have a musical lexicon (Ayotte et al., 2002; Graves et al., 2019; Tillmann et al., 2014), this kind of information in long-term memory about familiar melodies could influence potentially also short-term memory performance.

In contrast to performance level (as measured by *d'*), there was no effect of tonality on response bias *c*. While amusic participants showed a positive response bias *c* (a tendency to respond "same" for tonal and atonal sequences), control participants did not show a response bias in the present experiment, with *c* being close to 0. This observation is different from Albouy et al. (2013) reporting a positive response bias for both amusic and control participants. Interestingly, even for amusic participants, the response bias was weaker in the present experiment than in Albouy et al. (2013) study (with $c = 0.36$ and $c = 0.77$ in the two studies, respectively (across conditions)). One might argue that the weaker response bias (for both participant groups) might be related to the use of more complex material with longer duration. This interpretation seems in agreement with findings of Schulze et al. (2012) for non-musician participants: In contrast to Schulze et al.'s shorter sequences, which led to a positive response bias, the longest sequences (sequences of seven tones) led to a response bias *c* approaching 0 for tonal sequences, and participants were thus bias free.

The third dependent variable was correct response times. Here we did not observe an effect of tonality: participants responded as fast for tonal items as for atonal items. This finding differed from Albouy et al. (2013) where response times were faster for tonal trials than for atonal trials also for amusic participants. This difference might be related to the fact that we here measured response times from the start of the S2 item (leading to rather long response or decision times, 6724 ms on average), and participants were allowed responding before the end of S2, in contrast to Albouy et al (2013), measuring from the offset of S2 (with an average of 749 ms), and requesting participants to answer only after the end of S2. Nevertheless, response times in the present study revealed an interesting data pattern. For same trials, amusic and control participants responded similarly slowly, suggesting that they listened up to the end of S2 before answering. For different trials, however, amusic participants responded more slowly than did control participants, suggesting that amusic participants waited longer to get accumulating evidence for their answer than control participants who responded more rapidly when detecting the difference. This result might reflect amusics' increased uncertainty in perceiving or responding and their lower level of consciousness related to music and pitch processing. This observation is in agreement with other findings on congenital amusia, as for example related to participants' musical lexicon stored in long-term memory investigated with familiarity judgments in a gating paradigm (Tillmann et al., 2014). While amusic participants reached similar familiarity decisions as did control participants, the response time data revealed differences between amusic and control participants. For longer segments presented in the gating paradigm, amusics responded overall more slowly than did controls (i.e., for both familiar and unfamiliar excerpts). For shorter segments, amusics needed more time to reach their judgments for familiar excerpts, while they seemed to respond with the same speed as controls for unfamiliar excerpts. Similarly to our interpretation for the difference in response times for same vs. different trials, these results suggest amusics' need for additional processing because of their uncertainty and/or low confidence in their abilities. These results can be related with other recent findings that have led to the hypothesis that congenital amusia might be related to a disorder of conscious access to music processing rather than music processing per se (e.g., Omigie et al., 2012; Peretz et al., 2009; Stewart, 2011; Zendel et al., 2015). In terms of remediation strategies, this points to the interest to explore learning strategies that build on spared implicit processing. Methods such as errorless learning and vanishing cues tested in amnestic patients (e.g., Bier et al., 2002; Kessels & de Haan, 2003) could be considered also in congenital amusia.

In Albouy et al. (2013), we interpreted the observed tonality benefit on RTs as an indicator that amusic individuals still have some implicit processing of tonality, which might influence short-term memory performance. The findings of Albouy et al. (2013), which showed the tonality benefit only for response times, had led to the hypothesis that "amusics' short-term memory abilities might be improved for more complex tonal material (e.g., chord sequences or harmonized melodies, in contrast to single melodic lines as used in the present study)" (page 229), and that with the use of real musical material as in Tillmann et al. (2016) "increased short-term memory accuracy for tonal sequences might also be observed in amusics" (page 229). Our present study directly followed this suggestion, adapted and implemented the richer musical material in a short-term memory paradigm, and the findings confirmed the proposed hypothesis.

With the present behavioral paradigm, the findings cannot inform us whether the tonal benefit acts on the encoding, maintenance, and/or recall stages of short-term memory. Future research needs to use the present material and paradigm with joint neurophysiological recordings, aiming to reveal underlying neural correlates of each of the three stages involved in memory. Previous findings have shown that amusics' short-term memory deficit involves an altered fronto-temporal network, mostly in the right hemisphere (e.g., Albouy, Caclin, et al., 2019; Albouy et al., 2015; Albouy, Mattout, et al., 2013; Albouy, Peretz, et al., 2019), with deficits observed in all of the three memory stages listed above. In these studies, the short-term memory task was implemented with short isochronous tone sequences with a maximum length of six tones. These tones all belonged to the same tonality and are thus comparable to the tonal material used in Schulze et al. (2012) and Albouy et al. (2013). With this kind of material, the tonal benefit was not observed in amusics' memory performance level, but only on response times (Albouy, et al., 2013). It would thus be interesting to use our here introduced material in combination with electrophysiological or fMRI recordings to further show the neural correlates of tonal structure processing in short-term memory in non-amusic individuals as well as in interaction with amusics' deficits and their altered fronto-temporal network. The fMRI study by Schulze, Müller et al. (2011) pinpointed a difference between brain responses to tonal and atonal music in musicians, with reduced activation of the right IFG during encoding of tonal compared to atonal sequences (see also Cheung et al., 2018). This was supported by better behavioral memory performance for tonal sequences, while for nonmusicians, the task turned out to be overall too difficult, not revealing tonal vs. atonal differences for behavioral and neural data. Although difficult, the tasks we used in the present study and in Albouy et al. (2013) were do-able for the majority of our participants and enabled to observe a tonal advantage also for non-musicians. Amusia offers a new window to further understand auditory (in particular musical) working memory and its neural correlates (see Albouy, Peretz, et al., 2019), in particular to uncover the links between working memory and structure (tonality) processing networks. As the link between right frontal and temporal areas is impaired in amusia, music processing and encoding is likely to more strongly rely on the IPL, as a compensatory mechanism. Structural data from DTI tractography analyzed by Zhao et al. (2016), using graph theory, supported this hypothesis by showing an enhanced nodal strength in the right IPL in amusics relative to controls. As the arcuate fasciculus is affected in amusia, frontal and temporal regions can exchange information through more indirect routes connecting to the right IPL. This is also in line with the observation that fibers project from the right posterior Superior Temporal Gyrus (pSTG) toward the right IFG in controls, but toward the IPL in amusics (Loui et al., 2009). The functional magnetoencephalography study of Albouy et al. (2013) also found an increased gamma synchronization in the right temporo-parietal junction during maintenance of musical information in memory in amusics compared to controls, suggesting that amusics need to recruit a more distributed fronto-parietal network to perform the task. In the present study, we can hypothesize that the tonal structure of music might increase the involvement of the IPL (Schulze et al., 2011) and thus constitutes an effective support for musical memory in amusics. Future research could directly test this hypothesis, studying connectivity within the temporo-parieto-frontal network in amusics and control participants for the perception and memory of tonal and atonal musical materials.

CRediT authorship contribution statement

Yohana Lévêque: Investigation, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Philippe Lalitte:** Resources, Writing – review & editing. **Lesly Fornoni:** Investigation, Formal analysis, Data curation, Writing – review & editing. **Agathe Pralus:** Investigation, Writing – review & editing. **Philippe Albouy:** Conceptualization, Formal analysis, Validation, Writing – review & editing. **Patrick Bouchet:** Software, Writing – review & editing. **Anne Caclin:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision, Project administration. **Barbara Tillmann:** Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Celine Tourlonnias and Perrine Teyssier for help in running participants for the material test. This work was supported by a grant from “Agence Nationale de la Recherche” (ANR) of the French Ministry of Research ANR-11-BSH2-001-01 to BT and AC. This work was conducted in the framework of the LabEx CeLyA (“Centre Lyonnais d’Acoustique”, ANR-10-LABX-0060) and of the LabEx Cortex (“Construction, Function and Cognitive Function and Rehabilitation of the Cortex”, ANR-11-LABX-0042) of Université de Lyon, within the program “Investissements d’avenir” (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

Appendix A. Supplementary material

Sound files for the excerpts displayed in Figure 1.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandc.2022.105881.

References

Albouy, P., Caclin, A., Norman-Haignere, S. V., Lévêque, Y., Peretz, I., Tillmann, B., & Zatorre, R. J. (2019a). Decoding Task-Related Functional Brain Imaging Data to Identify Developmental Disorders: The Case of Congenital Amusia. *Frontiers in Neuroscience, 13*. https://doi.org/10.3389/fnins.2019.01165

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P.-E., … Tillmann, B. (2013a). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain: A. Journal of Neurology, 136*(Pt 5), 1639–1661. https://doi.org/10.1093/brain/awt082

Albouy, P., Mattout, J., Sanchez, G., Tillmann, B., & Caclin, A. (2015). Altered retrieval of melodic information in congenital amusia: Insights from dynamic causal modeling of MEG data. *Frontiers in Human Neuroscience, 9*, 20. https://doi.org/10.3389/fnhum.2015.00020

Albouy, P., Peretz, I., Bermudez, P., Zatorre, R. J., Tillmann, B., & Caclin, A. (2019b). Specialized neural dynamics for verbal and tonal memory: FMRI evidence in congenital amusia. *Human Brain Mapping, 40*(3), 855–867. https://doi.org/10.1002/hbm.24416

Albouy, P., Schulze, K., Caclin, A., & Tillmann, B. (2013b). Does tonality boost short-term memory in congenital amusia? *Brain Research, 1537*, 224–232. https://doi.org/10.1016/j.brainres.2013.09.003

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusiaA group study of adults afflicted with a music-specific disorder. *Brain, 125*(2), 238–251. https://doi.org/10.1093/brain/awf028

Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 21*(6), 1278–1296. https://doi.org/10.1037/0096-1523.21.6.1278

Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5*(1), 1–30. https://doi.org/10.2307/40285384

Bharucha, J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13*(1), 63–102. https://doi.org/10.1016/0010-0277(83)90003-3

Bier, N., Vanier, M., & Meulemans, T. (2002). Errorless learning: A method to help amnesic patients learn new information. *Journal of Cognitive Rehabilitation, 20*. https://orbi.uliege.be/handle/2268/1951.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*(1), 100–130. https://doi.org/10.1016/j.cognition.2005.11.007

Bor, D., Cumming, N., Scott, C. E. L., & Owen, A. M. (2004). Prefrontal cortical involvement in verbal encoding strategies. *The European Journal of Neuroscience, 19*(12), 3365–3370. https://doi.org/10.1111/j.1460-9568.2004.03438.x

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron, 37*(2), 361–367. https://doi.org/10.1016/s0896-6273(02)01171-6

Botvinick, M., & Bylsma, L. M. (2005). Regularization in short-term memory for serial order. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(2), 351–358. https://doi.org/10.1037/0278-7393.31.2.351

Cheung, V. K. M., Meyer, L., Friederici, A. D., & Koelsch, S. (2018). The right inferior frontal gyrus processes nested non-local dependencies in music. *Scientific Reports, 8*(1), 3822. https://doi.org/10.1038/s41598-018-22144-9

Cousineau, M., McDermott, J. H., & Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. *Proceedings of the National Academy of Sciences of the United States of America, 109*(48), 19858–19863. https://doi.org/10.1073/pnas.1207989109

Cowan, N. (1999). The differential maturation of two processing rates related to digit span. *Journal of Experimental Child Psychology, 72*(3), 193–209. https://doi.org/10.1006/jecp.1998.2486

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169*, 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics, 50*(4), 305–313. https://doi.org/10.3758/bf03212222

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Filipic, S., Tillmann, B., & Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review, 17*(3), 335–341. https://doi.org/10.3758/PBR.17.3.335

Foote, J. (2000). Automatic audio segmentation using a measure of audio novelty. *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532), 1*, 452–455 vol.1. https://doi.org/10.1109/ICME.2000.869637.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying ‘tone deafness’. *Brain, 127*(4), 801–810. https://doi.org/10.1093/brain/awh105

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., & Schlaug, G. (2003). Functional anatomy of pitch memory—An fMRI study with sparse temporal sampling. *NeuroImage, 19*(4), 1417–1426.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P.-C.-H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5*(6), 236–243. https://doi.org/10.1016/s1364-6613(00)01662-4

Gosselin, N., Jolicoeur, P., & Peretz, I. (2009). Impaired memory for pitch in congenital amusia. *Annals of the New York Academy of Sciences, 1169*, 270–272. https://doi.org/10.1111/j.1749-6632.2009.04762.x

Gosselin, N., Paquette, S., & Peretz, I. (2015). Sensitivity to musical emotions in congenital amusia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 71*, 171–182. https://doi.org/10.1016/j.cortex.2015.06.022

Graves, J. E., Pralus, A., Fornoni, L., Oxenham, A. J., Caclin, A., & Tillmann, B. (2019). Short- and long-term memory for pitch and non-pitch contours: Insights from congenital amusia. *Brain and Cognition, 136*, Article 103614. https://doi.org/10.1016/j.bandc.2019.103614

Griffiths, T. D. (2001). The neural processing of complex sounds. *Annals of the New York Academy of Sciences, 930*, 133–142. https://doi.org/10.1111/j.1749-6632.2001.tb05729.x

Griffiths, T. D., & Green, G. G. R. (1999). Cortical Activation during Perception of a Rotating Wide-Field Acoustic Stimulus. *NeuroImage, 10*(1), 84–90. https://doi.org/10.1006/nimg.1999.0464

Guo, S., & Koelsch, S. (2016). Effects of veridical expectations on syntax processing in music: Event-related potential evidence. *Scientific Reports, 6*(1), 19064. https://doi.org/10.1038/srep19064

Harte, C., Sandler, M., & Gasser, M. (2006). Detecting harmonic change in musical audio. *Proceedings of the 1st ACM Workshop on Audio and Music Computing Multimedia - AMCMM '06*, 21. https://doi.org/10.1145/1178723.1178727.

Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical Thickness in Congenital Amusia: When Less Is Better Than More. *Journal of Neuroscience, 27*(47), 13028–13032. https://doi.org/10.1523/JNEUROSCI.3039-07.2007

Hyde, K. L., & Peretz, I. (2004). Brains That Are out of Tune but in Time. *Psychological Science, 15*(5), 356–360. https://doi.org/10.1111/j.0956-7976.2004.00683.x

Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: A two-site study. *Brain: A. Journal of Neurology, 129*(Pt 10), 2562–2570. https://doi.org/10.1093/brain/awl204

Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex (New York, N.Y.: 1991), 21*(2), 292–299. https://doi.org/10.1093/cercor/bhq094.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience, 6*(7), 682–687. https://doi.org/10.1038/nn1081

Jones, G., & Macken, B. (2018). Long-term associative learning predicts verbal short-term memory performance. *Memory & Cognition, 46*(2), 216–229. https://doi.org/10.3758/s13421-017-0759-3

Jones, G., Tamburelli, M., Watson, S. E., Gobet, F., & Pine, J. M. (2010). Lexicality and frequency in specific language impairment: Accuracy and error data from two nonword repetition tests. *Journal of Speech, Language, and Hearing Research: JSLHR, 53*(6), 1642–1655. https://doi.org/10.1044/1092-4388(2010/09-0222)

Justus, T. C., & Bharucha, J. J. (2001). Modularity in musical processing: The automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance, 27*(4), 1000–1011. https://doi.org/10.1037/0096-1523.27.4.1000

Kessels, R. P. C., & de Haan, E. H. F. (2003). Implicit learning in memory rehabilitation: A meta-analysis on errorless learning and vanishing cues methods. *Journal of Clinical and Experimental Neuropsychology, 25*(6), 805–814. https://doi.org/10.1076/jcen.25.6.805.16474

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage, 25*(4), 1068–1076. https://doi.org/10.1016/j.neuroimage.2004.12.050

Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: An ERP study. *Brain Research, 1212*, 55–62. https://doi.org/10.1016/j.brainres.2007.10.078

Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants' Perception of Phrase Structure in Music. *Psychological Science, 1*(1), 70–73. https://doi.org/10.1111/j.1467-9280.1990.tb00070.x

Kumar, S., Joseph, S., Gander, P. E., Barascud, N., Halpern, A. R., & Griffiths, T. D. (2016). A Brain System for Auditory Working Memory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 36*(16), 4492–4505. https://doi.org/10.1523/JNEUROSCI.4341-14.2016

Lalitte, P., Bigand, E., Kantor-Martynuska, J., & Delbéé, C. (2009). On Listening to Atonal Variants of Two Piano Sonatas by Beethoven. *Music Perception, 26*(3), 223–234. https://doi.org/10.1525/mp.2009.26.3.223

Lartillot, O., & Toiviainen, P. (2007). Mir in Matlab (ii): A Toolbox for Musical Feature Extraction from Audio. *Austrian Computer Society*.

Lerch, A. (2012). *An Introduction to Audio Content Analysis: Applications in Signal Processing and Music Informatics*. John Wiley & Sons.

Leveque, Y., Fauvel, B., Groussard, M., Caclin, A., Albouy, P., Platel, H., & Tillmann, B. (2016). Altered intrinsic connectivity of the auditory cortex in congenital amusia. *Journal of Neurophysiology*, jn.00663.2015. https://doi.org/10.1152/jn.00663.2015.

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain: A. Journal of Neurology, 133*(Pt 6), 1682–1693. https://doi.org/10.1093/brain/awq089

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 29*(33), 10215–10220. https://doi.org/10.1523/JNEUROSCI.1701-09.2009

Majerus, S., Martinez Perez, T., & Oberauer, K. (2012). Two distinct origins of long-term learning effects in verbal short-term memory. *Journal of Memory and Language, 66*(1), 38–51. https://doi.org/10.1016/j.jml.2011.07.006

Marin, M. M., Gingras, B., & Stewart, L. (2012). Perception of musical timbre in congenital amusia: Categorization, discrimination and short-term memory. *Neuropsychologia, 50*(3), 367–378. https://doi.org/10.1016/j.neuropsychologia.2011.12.006

Marin, M. M., Thompson, W. F., Gingras, B., & Stewart, L. (2015). Affective evaluation of simultaneous tone combinations in congenital amusia. *Neuropsychologia, 78*, 207–220. https://doi.org/10.1016/j.neuropsychologia.2015.10.004

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158

Omigie, D., Müllensiefen, D., & Stewart, L. (2012). The Experience of Music in Congenital Amusia. *Music Perception: An Interdisciplinary Journal, 30*(1), 1–18. https://doi.org/10.1525/mp.2012.30.1.1

Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia, 51*(9), 1749–1762. https://doi.org/10.1016/j.neuropsychologia.2013.05.010

Omigie, D., & Stewart, L. (2011). Preserved statistical learning of tonal and linguistic material in congenital amusia. *Frontiers in Psychology, 2*, 109. https://doi.org/10.3389/fpsyg.2011.00109

Peretz, I. (2016). Neurobiology of Congenital Amusia. *Trends in Cognitive Sciences, 20*(11), 857–867. https://doi.org/10.1016/j.tics.2016.09.002

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33*(2), 185–191. https://doi.org/10.1016/s0896-6273(01)00580-3

Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain: A. Journal of Neurology, 132*(Pt 5), 1277–1286. https://doi.org/10.1093/brain/awp055

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999*, 58–75.

Peretz, I., Saffran, J., Schön, D., & Gosselin, N. (2012). Statistical learning of speech, not music, in congenital amusia. *Annals of the New York Academy of Sciences, 1252*, 361–367. https://doi.org/10.1111/j.1749-6632.2011.06429.x

Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S., … Eustache, F. (1997). The structural components of music perception. A functional anatomical study. *Brain: A. Journal of Neurology, 120*(Pt 2), 229–243. https://doi.org/10.1093/brain/120.2.229

Plomp, R., & Levelt, W. J. (1965). Tonal consonance and critical bandwidth. *The Journal of the Acoustical Society of America, 38*(4), 548–560. https://doi.org/10.1121/1.1909741

Quiroga-Martinez, D. R., Tillmann, B., Brattico, E., Cholvy, F., Fornoni, L., Vuust, P., & Caclin, A. (2021). Listeners with congenital amusia are sensitive to context uncertainty in melodic sequences. *Neuropsychologia, 107911*. https://doi.org/10.1016/j.neuropsychologia.2021.107911

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences, 11*(5), 211–218. https://doi.org/10.1016/j.tics.2007.02.006

Schubotz, R. I., von Cramon, D. Y., & Lohmann, G. (2003). Auditory what, where, and when: A sensory somatotopy in lateral premotor cortex. *NeuroImage, 20*(1), 173–185. https://doi.org/10.1016/s1053-8119(03)00218-0

Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences, 1252*, 229–236. https://doi.org/10.1111/j.1749-6632.2012.06447.x

Schulze, K., Mueller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *European Journal of Neuroscience, 33*(1), 189–196. https://doi.org/10.1111/j.1460-9568.2010.07470.x

Stevens, A. A. (2004). Dissociating the cortical basis of memory for voices, words and tones. *Brain Research. Cognitive Brain Research, 18*(2), 162–171. https://doi.org/10.1016/j.cogbrainres.2003.10.008

Stewart, L. (2011). Characterizing congenital amusia. *Quarterly Journal of Experimental Psychology (2006), 64*(4), 625–638. https://doi.org/10.1080/17470218.2011.552730.

Tillmann, B., Albouy, P., & Caclin, A. (2015). Congenital amusias. *Handbook of Clinical Neurology, 129*, 589–605. https://doi.org/10.1016/B978-0-444-62630-1.00033-0

Tillmann, B., Albouy, P., Caclin, A., & Bigand, E. (2014). Musical familiarity in congenital amusia: Evidence from a gating paradigm. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 59*, 84–94. https://doi.org/10.1016/j.cortex.2014.07.012

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*(4), 885–913. https://doi.org/10.1037/0033-295x.107.4.885

Tillmann, B., Gosselin, N., Bigand, E., & Peretz, I. (2012). Priming paradigm reveals harmonic structure processing in congenital amusia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 48*(8), 1073–1078. https://doi.org/10.1016/j.cortex.2012.01.001

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16*(2), 145–161. https://doi.org/10.1016/S0926-6410(02)00245-8

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., & von Cramon, D. Y. (2006). Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *NeuroImage, 31*(4), 1771–1782. https://doi.org/10.1016/j.neuroimage.2006.02.028

Tillmann, B., Lalitte, P., Albouy, P., Caclin, A., & Bigand, E. (2016a). Discrimination of tonal and atonal music in congenital amusia: The advantage of implicit tasks. *Neuropsychologia, 85*, 10–18. https://doi.org/10.1016/j.neuropsychologia.2016.02.027

Tillmann, B., Lévêque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. *Brain Research, 1640, Part B*, 251–263. https://doi.org/10.1016/j.brainres.2015.10.035.

Tillmann, B., Schulze, K., & Foxton, J. M. (2009). Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition, 71*(3), 259–264. https://doi.org/10.1016/j.bandc.2009.08.003

Tulving, E., & Patkau, J. E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology, 16*, 83–95. https://doi.org/10.1037/h0083231

Weiss, M. W., & Peretz, I. (2019). Ability to process musical pitch is unrelated to the memory advantage for vocal music. *Brain and Cognition, 129*, 35–39. https://doi.org/10.1016/j.bandc.2018.11.011

Williamson, V. J., McDonald, C., Deutsch, D., Griffiths, T., & Stewart, L. (2010). Faster decline of pitch memory over time in congenital amusia. *Advances in Cognitive Psychology, 6*, 15–22. https://doi.org/10.2478/v10053-008-0073-5

Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem. *Memory (Hove, England), 18*(6), 657–669. https://doi.org/10.1080/09658211.2010.501339

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 14*(4), 1908–1919.

Zendel, B. R., Lagrois, M.-É., Robitaille, N., & Peretz, I. (2015). Attending to Pitch Information Inhibits Processing of Pitch Information: The Curious Case of Amusia. *The Journal of Neuroscience, 35*(9), 3815–3824. https://doi.org/10.1523/JNEUROSCI.3766-14.2015

Zhao, Y., Chen, X., Zhong, S., Cui, Z., Gong, G., Dong, Q., & Nan, Y. (2016). Abnormal topological organization of the white matter network in Mandarin speakers with congenital amusia. *Scientific Reports, 6*(1), 26505. https://doi.org/10.1038/srep26505