Brain and Cognition 71 (2009) 259–264

Contents lists available at ScienceDirect

Brain and Cognition

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

Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds

Barbara Tillmann [a,*], Katrin Schulze [a,b], Jessica M. Foxton [c,d]

[a] Université Claude Bernard Lyon 1, Neurosciences Sensorielles Comportement Cognition CNRS-UMR 5020, IFR 19, Lyon, France
[b] University College of London, Developmental Cognitive Neuroscience Unit, Institute of Child Health, London, UK
[c] INSERM U821, IFR 19, Lyon, France
[d] Cerco, CNRS-UMR 5549, Université Paul Sabatier, Toulouse, France

ARTICLE INFO

Article history:
Accepted 7 August 2009
Available online 16 September 2009

Keywords:
Congenital amusia
Short-term memory
Pitch
Timbre
Words
Contour
Auditory scene analysis

ABSTRACT

Congenital amusia refers to a lifelong disorder of music processing and is linked to pitch-processing deficits. The present study investigated congenital amusics' short-term memory for tones, musical timbres and words. Sequences of five events (tones, timbres or words) were presented in pairs and participants had to indicate whether the sequences were the same or different. The performance of congenital amusics confirmed a memory deficit for tone sequences, but showed normal performance for word sequences. For timbre sequences, amusics' memory performance was impaired in comparison to matched controls. Overall timbre performance was found to be correlated with melodic contour processing (as assessed by the Montreal Battery of Evaluation of Amusia). The present findings show that amusics' deficits extend to non-verbal sound material other than pitch, in this case timbre, while not affecting memory for verbal material. This is in line with previous suggestions about the domain-specificity of congenital amusia.

1. Introduction

Congenital amusia (also named tone or tune deafness) refers to a lifelong disorder of music processing that occurs without brain damage, and is estimated to affect about 4% of the general population (see Peretz & Hyde, 2003 for a review). Individuals with congenital amusia have difficulty recognizing familiar tunes without lyrics and detecting a wrong or out-of-tune note. The musical disorder occurs despite normal performance on tests of intelligence, auditory processing, cognitive functioning, language processing, and verbal memory skills (see Ayotte, Peretz, & Hyde, 2002; Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Peretz et al., 2002, for extensive testing).

This disorder has been recognised for a long time, but has only been systematically studied relatively recently (see Ayotte et al., 2002), mainly thanks to the development of the Montreal Battery for the Evaluation of Amusia MBEA (Peretz, Champod, & Hyde, 2003). In the MBEA, seven sub-tests address various components of music perception and memory, notably the pitch dimension (detection of an out-of-key note, a contour violation, or interval changes), the time dimension (rhythm and meter perception) and incidental memory (i.e., for melodies used in preceding sub-tests). Although some amusic individuals show deficits for rhythm perception, the major deficit concerns the pitch dimension (as assessed by the MBEA scale, contour and interval sub-tests).

The pitch deficit is not limited to musical contexts, it also affects basic pitch discrimination in unmusical tone material. Performance is impaired for the recognition of pitch direction, the perception of more complex pitch patterns and the detection of pitch changes in continuous and discrete sounds as well as in isochronous sequences (Foxton et al., 2004; Hyde & Peretz, 2004). Amusic individuals have difficulty detecting pitch changes smaller than two semitones (or even more, see Peretz et al., 2002). In earlier studies, normal performance has been reported for the processing of intonation in speech (Ayotte et al., 2002; Patel, Foxton, & Griffiths, 2005; Peretz et al., 2002), thus implying that the pitch deficit only affects music perception. Recent data suggest that for some amusics, the processing advantage for large pitch differences in speech might not be observed for gliding pitch changes at a slow rate (Patel, Wong, Foxton, Lochy, & Peretz, 2008).

The tasks revealing pitch deficits involved not only perception, but also memory components: Two melodies, tone pairs or tones have to be compared (same-different paradigms) or the odd-one out has to be found among three events (i.e., AAX versus AXA). It has been argued that memory and/or attention deficits cannot explain the poor performance because amusic individuals perform

* Corresponding author. Address: Université Claude Bernard Lyon I, CNRS UMR 5020 Neurosciences Comportement Cognition, 50 Av. Tony Garnier, F-69366 Lyon Cedex 07, France. Fax: +33 (0) 4 37 28 76 01.
E-mail address: barbara.tillmann@olfac.univ-lyon1.fr (B. Tillmann).

0278-2626/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandc.2009.08.003

well for large pitch differences and show normal performance on standard memory tasks, such as the digit span from the Wechsler Adult Intelligence Scale III. However, another possibility might be that verbal and non-verbal memory differ (as suggested, for example, by Deutsch, 1970), with only the latter being impaired in congenital amusia. Recent research investigating pitch memory with variations in sequence length and intervening delay, provide direct evidence for short-term memory deficits for pitch in congenital amusia (Gosselin, Jolicoeur, & Peretz, 2008; Stewart, McDonald, Kumar, Deutsch, & Griffiths, 2008). However, up to now, no study has investigated memory performance for non-verbal material other than pitch. For non-verbal sound material, such as voices (i.e., of famous individuals) and environmental sounds (i.e., animal sounds, industrial sounds, human noises), Peretz et al. (2002) and Ayotte et al. (2002) have only shown that individuals with congenital amusia show normal performance in naming and recognition tasks (testing for access to knowledge stored in long-term memory, without testing for short-term memory).

Our present study investigated short-term memory for tones, musical timbres and words. Previous research using interference memory paradigms suggests the existence of short-term memory modules that are specialized for the retention of either pitch (without storing other attributes, Semal & Demany, 1991; Semal, Demany, Ueda, & Hallé, 1996) or timbre (Starr & Pitt, 1997). Timbre enables listeners to distinguish between different instruments or speakers, and plays a role in sequence perception as well as the separation of sound sources (Bregman, 1990). The Acoustical Society defines timbre with reference to the features that enable the distinction between two sounds of identical pitch, intensity, duration and location. Research has shown that timbre is a multidimensional set of auditory attributes based on the temporal and spectral features of sounds (cf. Krumhansl, 1989; McAdams, Winsberg, Donnadieu, DeSoete, & Krimphoff, 1995; Samson, Zatorre, & Ramsay, 1997). Models of timbre have been based on perceived similarity judgments and propose mental representations of timbre in three-dimensional spatial structures with axes representing attack time (the time to reach the maximum of the energy envelop), spectral centroid (or spectral center of gravity) and spectral flux (how the spectral envelope changes over time).

In the present study, participants (amusics and matched controls) listened to five-event sequences, with the events being either tones, words or timbres. The tones were played by the same instrument, but differed in pitch. The words were spoken by the same voice at the same pitch, but differed in the phonemes and semantic content. The timbres were played at the same pitch, but differed in spectro-temporal information. The five-event sequences were presented in pairs and were separated by a 3-s silent delay. They were either the same or differed in the order of presentation of the events (i.e., two events were exchanged). For the pitch memory task, the differences between the tones were altered according to the amusic participants' pitch perception thresholds. Based on previous findings on congenital amusia, we expected a memory deficit for tone sequences, but normal performance for word sequences. It was reasoned that if the memory deficit is restricted to pitch, performance should be normal for the timbre sequences. If, however, the deficit more generally affects memory for non-verbal sounds, impaired performance should be observed for the timbre sequences. The present paradigm has previously been tested on healthy students with varying sequence lengths (Schulze & Tillmann, 2007), and we expected matched control groups to replicate the students' data for the five-event sequences.[1] With the aim of investigating potential memory deficits in amusic listeners, the present study focused on comparisons between amusic and control participants for each of the materials.

2. Method

2.1. Participants

The amusic group consisted of 10 adults (six women) with a mean age of 33.6 years (SD = 10.3). Their level of education was on average 15.2 years (SD = 1.87), and the average musical training was 0.85 years (SD = 1.67). The matched control group consisted of 10 adults (six women) with a mean age of 36.2 years (SD = 10.3), a reported level of education of 13.9 years (SD = 1.56) and an average musical training of 0.3 years (SD = .95). All participants performed the MBEA: The average score for the amusic group (21.07; SD = 1.59) differed significantly from the score for the control group (27.23; SD = 1.18), $t(18) = 9.85$, $p < .0001$. All amusics obtained scores below the cut-off score (23, which is two standard deviations below the norm), except for one participant who obtained a score of 23.67. This amusic was included because of low performance on the scale sub-test (18, with a cut-off at 22), which has been shown to be strongly diagnostic (Peretz et al., 2008). All controls performed above the cut-off score (see Table 1 of Supplementary material).

2.2. Pretest: pitch difference detection thresholds

2.2.1. Method and procedure

To determine perceptual thresholds for detecting differences in pitch, a two-alternative forced choice task was employed. Participants were presented with two pairs of pure tones: One pair of tones that had the same pitch, and a second pair of tones that differed in pitch. Participants were asked to decide whether the first or the second pair contained the pitch difference. Adaptive tracking using a two-down, one-up staircase procedure was employed to determine the perceptual threshold, targeting 70.7% correct performance. The sound pairs consisted of two 100 ms pure tones, gated with 10 ms amplitude ramps, and separated by a silent interval of 150 ms. The frequency of the same-pitch pairs was 512 Hz; the different-pitch pairs always contained one tone of 512 Hz and another tone of a higher frequency that was randomly presented first or second within the pair. During the task, there was no time limit for the response, and the next sound pair was presented 650 ms after the response. The order of the same-pitch and different-pitch pairs was randomized. The test was administered in three runs of 30 trials, and the pitch difference at the beginning of the first run was set at 2.0 semitones. The step size in the adaptive track was 0.1 semitones. Thresholds were determined by averaging the last six changes in the direction of the adaptive track. Prior to the task, participants completed a short practice run of six items, where they were given error feedback. No feedback was given during the three adaptive runs.

2.2.2. Results

Thresholds for the amusic group ranged from 0.2 to 4 semitones (mean = 1.32, SD = 1.17) and for the control group from .07 to 1.67 (mean = 0.57, SD = .61) (see Table 1 of Supplementary material). This difference was statistically significant using a one-tailed *t*-test, $t(18) = 1.80$, $p = .045$. For the amusics, five participants had thresholds below one semitone (average threshold of .54), three had thresholds above one semitone (average threshold of 1.28) and two above 2 semitones (thresholds of 2.6 and 4, respectively). For the controls, seven participants had thresholds below one semitone (average threshold of 0.22), and three participants had thresholds above one semitone (with respectively 1.2, 1.3 and 1.7 semitones). For the amusics only, these results were used to define the tone sequences of the short-term memory task, as detailed be-

[1] For the five-event sequences, students' performance level for timbres was as good as for words (i.e., Hits – False Alarms of .43 and .45, respectively), while performance was best for tones (i.e., Hits – False Alarms of .61).

low. The observed overlap in pitch thresholds between amusic and control groups is in agreement with previous findings for thresholds (Foxton et al., 2004), melody familiarity judgments and recognition (Ayotte et al., 2002) and congruity/incongruity judgments of out-of-key or out-of-tune tones in melodies (Peretz, Brattico, Järvenpää, & Tervaniemi, 2009).

2.3. Short-term memory task for tones, timbres and words

2.3.1. Material and apparatus

The short-term memory task was constructed in the same way for each condition: Participants listened to an auditory sequence, which consisted of five events (tones, timbres or words), followed by silence for 3 s, and then a second sequence, in which the five events were played in either the same or a different order. If the order was different, two events within the sequence were exchanged (e.g., A B C D E – A D C B E). This manipulation never involved the first element of the sequence. Each event (i.e., pitch, timbre, word) had a duration of 500 ms and the five events in the sequence were presented with an inter-stimulus-interval of 40 ms. The software Presentation (Neurobehavioral systems) was used to run the experiment and record the responses. Sound examples are available as Supplementary material.

2.3.2. Pitch condition

Three sets of six piano tones differing in pitch height were used to create the five-event sequences. The sets differed in the size of the intervals between the tones, but all used tones that belonged to the key of C Major. The note E3 was used as the referent pitch, occurring in all of the sets, and tones above and below E3 were changed to vary the interval sizes. In Set 1 (with intervals of 1 and 2 semitones between adjacent tones in the set), the notes C3, D3, E3, F3, G3 and A3 were used (average interval size of 1.8 semitones). For Sets 2 and 3, the intervals between the tones were stretched by a factor of 2 and 3, respectively, with the constraint that the new set remained tonal (in the key of C Major). For example, for Set 2, an interval of two semitones in Set 1 was stretched to four semitones and to avoid atonality, this interval might be adjusted (by one semitone above or below). In Set 2, the notes A2, C3, E3, G3, C4 and E4 were used; here the smallest interval size was 3 semitones, with an average interval of 3.8 (ranging from 3 to 5 semitones). In Set 3, the notes E2, A2, E3, G3, C4 and G4 were used; here the smallest interval size was 3, with an average interval size of 5.4 (ranging from 3 to 7 semitones). Set 1 was used for all of the controls. For the amusics, the set varied according to each individual's pitch threshold: Sets 1, 2 and 3 were used, respectively, for amusics with thresholds below one semitone, below 2 semitones and above 2 semitones.[2] The stimuli were created with the software Digital performer, Cubase 5.1 (Steinberg) and a Halion Sampler (Steinberg) using an acoustic piano timbre. For the 'different' trials, the exchanged tones preserved the contour of the sequence in 11 of the 14 trials and violated the contour in the other 3 trials. The contour-violation was a consequence of the constraints used in constructing the sequences: namely the use of six events, the control of the frequency of occurrence of each event over all sequences and in each of the five positions, and finally the creation of 'different' trials by exchanging events (except for the first) without introducing new events.

2.3.3. Timbre condition

Six timbres were used (guitar, cello, flute, trumpet, vibes, piano), all played at the same pitch at 330 Hz (i.e., E3). The stimuli were created with the software Cubase 5.1 (Steinberg) and a Halion Sampler (Steinberg). The loudness of the timbres was matched subjectively and adjusted with Adobe Audition digital sound software.

2.3.4. Word condition

Six monosyllabic French words were used: toux (cough), loup (wolf), boue (dirt), mou (lung), goût (taste) and poux (bug), spoken by a female voice. All recordings were adjusted to the pitch of 230 Hz with STRAIGHT (Kawahara & Irino, 2004), and subjectively equalized in loudness using Adobe Audition digital sound software. The words were selected from a pool of recorded words on the basis of subjective ratings indicating easy intelligibility (i.e., using a subjective scale from 1 (very easy to understand) to 5 (not easy to understand)) by eight native French speakers. Phonologically similar monosyllabic words were chosen to increase the difficulty of the verbal task, as it is known that there is lower performance for phonologically similar consonant sequences than for phonological different-sounding consonant sequences (Conrad & Hull, 1964).

2.3.5. Procedure

For each condition, there were two blocks (i.e., two timbre, two word, and two pitch blocks). At the beginning of the first block of each condition, two example trials (one same, one different) were presented. Example and experimental trials consisted of a five-event sequence, a silence of 3 s and a second five-event sequence. Participants were required to listen carefully to the sequences and to indicate whether the second sequence was identical to the first or whether the order of the events had been changed. They were informed that a different trial was based on the same events, arranged in a different order. Participants indicated their answers by button presses. No feedback was given for the experimental trials. Half of the sequence pairs were the same, and the other half were different. For each condition, 28 trials were presented (14 same and 14 different pairs), resulting in 84 trials in total. The blocks were presented in one of six different orders for each participant (the exact order being randomly chosen). Within each block, the trials were presented in a pseudo-randomized order, with the constraint that a given sequence could not be presented on consecutive trials, and that the same trial type (i.e., same, different) could not be repeated more than three times in a row.

3. Results

Performance was analyzed by calculating Hits (number of correct responses for different trials/number of different trials) minus False Alarms FAs (number of incorrect responses for same trials/number of same trials) (Fig. 1 and Table 1 of Supplementary material): Performance in all conditions was significantly above chance, both for the controls ($p < .0001$ for all conditions) and for the amusics ($p = .0499$ for pitch; $p = .02$ for timbre; $p = .0001$ for words). Hits-FAs were analyzed with a 2×3 ANOVA with Group (amusics/controls) as the between-participants factor and Material (pitch, timbre, words) as the within-participants factor. The main effect of Group was significant, $F(1,18) = 27.32$, $MSE = .05$, $p < .0001$, as was the interaction between Group and Material, $F(2,36) = 9.06$, $MSE = .04$, $p = .0007$. Planned comparisons testing for group differences indicated that the performance of the amusic group was lower than the performance of the control group for pitch, $F(1,18) = 40.76$, $p < .001$, and for timbre, $F(1,18) = 5.17$, $p = .04$, but not for words, $p = .41$. Partial η^2 for these three between-group comparisons within each stimulus type were .69, .22 and .04, respectively.

For the pitch task, two additional analyses were performed to investigate two different features of the material. (1) The first analysis investigated whether the difference between amusic and con-

[2] We did not adopt sets stretched any wider than Set 3 in order to avoid extreme ranges between the highest and lowest tones (i.e., 36 semitones).

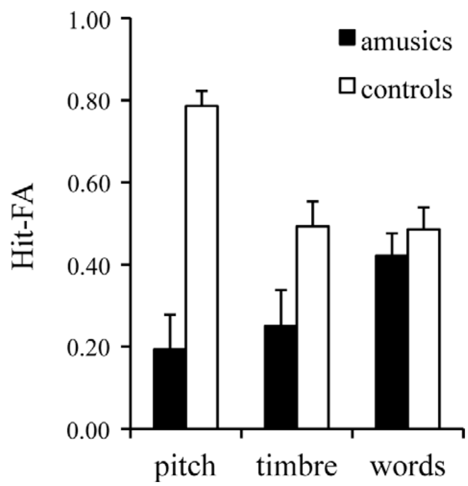

Fig. 1. Performance of the amusic and control groups in terms of Hits minus False Alarms (FA), presented as a function of the material used in the short-term memory task (pitch, timbre, words). Error bars represent the between-participants standard errors.

trol participants might have been influenced by the presence/absence of contour-violation in the different trials. We calculated the percentage of correct responses separately for the contour-violating and the contour-preserving trials ($n = 3$ and $n = 11$, respectively). While the contour-violating trials led to better performance than the contour-preserving trials for both groups, the amusic participants performed significantly below the control participants for both trial types (56.67% versus 96.67%, $t(18) = 3.80$, $p < .01$ for contour-violating trials; 39.09% versus 78.18%, $t(18) = 6.41$, $p < .01$ for contour-preserving trials). Thus, the amusics' pitch processing deficit was observed independently of the possibility of using melodic contour as a cue. Interestingly, the amusics' performance reflected some sensitivity to melodic contour cues, as did the controls, but this observation needs to be further investigated by systematically manipulating the presence/absence of contour violation with equal trial numbers. (2) The second analysis addressed the question of whether the amusics' weaker performance might have been an artifact resulting from the material construction, specifically, whether the deficit was caused by more pitch chroma repetition in the sequences with the larger pitch differences, which were only used for certain amusic participants and never for the controls. To investigate this possibility, we compared the performance of two amusic sub-groups: those having performed the task with the same pitch sequences as the controls, and those having performed the task with the stretched pitch sequences. We found that the former amusic sub-group obtained a mean score of .26 (SD = .26, ranging from −.14 to .57), compared to .13 (SD = .29, ranging from −.21 to .36) for the latter sub-group. Pitch performance thus considerably overlapped between these two sub-groups, and the individual performance of all of the amusics was below the performance of the control group (mean of .79, SD = .12, ranging from .64 to .92). Thus, the amusics' impaired pitch performance cannot be explained by differences in pitch chroma introduced by the experimental manipulations.

For each of the three conditions, we calculated correlations between performance levels and the scores on the MBEA. With the total MBEA score, the correlation was only significant for the pitch condition, $r(18) = .85$, $p < .0001$. When considering the MBEA subtest scores separately, the correlations were significant between the pitch condition and all of the sub-tests (scale: $r(18) = .80$, $p < .0001$; contour: $r(18) = .87$, $p < .0001$; interval: $r(18) = .74$, $p < .0001$; rhythm: $r(18) = .73$, $p < .0001$; meter: $r(18) = .54$, $p < .02$; memory: $r(18) = .56$, $p < .01$), and also between the timbre condition and the contour sub-test, $r(18) = .47$, $p < .04$ (Fig. 2). The correlation between the timbre condition and the contour sub-test was partly driven by the group differences: For amusics only, this correlation was positive, but not significant, $r(8) = .39$, $p = .27$, while it was absent for controls, $r(7) = .02$ (without the outlier). There were no significant correlations between the word condition and the MBEA sub-tests.

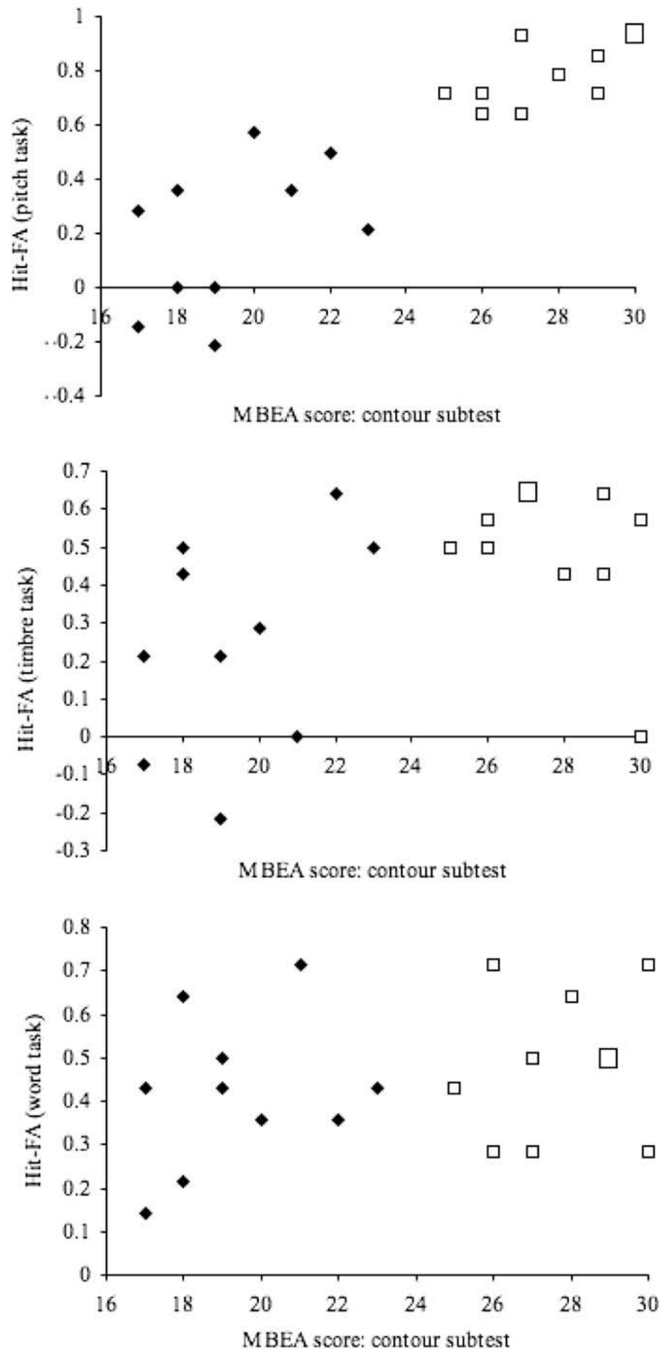

Fig. 2. Scatter plots for performance on the pitch, timbre and word tasks against the MBEA scores for the contour sub-test. Scores, expressed as Hits minus False Alarms (FA), are shown for both the amusic participants (black) and the control participants (white). The larger square represents two participants with identical scores.

Finally, it is worth noting that we observed a significant correlation between performance on the pitch and timbre conditions (Fig. 3), $r(18) = .59$, $p < .007$ (without the one control outlier, $r(18) = .75$, $p < .0001$). This correlation was due to the amusic

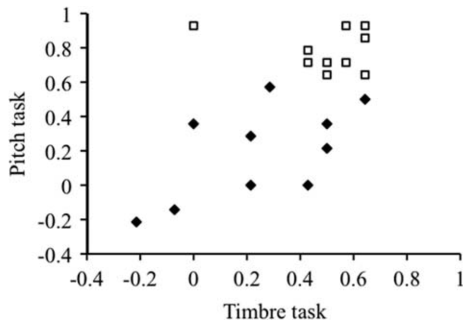

Fig. 3. Scatter plot for performance, expressed as Hits minus False Alarms (FA), on the pitch and timbre tasks for both the amusic participants (black) and the control participants (white).

group, $r(8) = .60$, $p < .07$ (but $p > .36$ for the control group, with and without outlier). The correlations between performance on the pitch and word conditions ($r(18) = 0.16$; $p = .51$) and between the timbre and word conditions ($r(18) = -.12$; $p = .63$), were not significant.

4. Discussion

The present study investigated short-term memory for pitch, timbre and words in congenital amusics and matched controls. The control group replicated our previous findings in students, showing similar performance for timbres and words, but better performance for pitch. In contrast to the controls, the amusics' performance was most strongly impaired for the pitch material, slightly – though significantly – impaired for the timbre material, but unimpaired for the verbal material. This finding suggests that the deficit in congenital amusia only affects short-term memory for non-verbal sound materials. This is in line with previous suggestions about the domain-specificity of the deficit (Peretz & Hyde, 2003). At first sight, this seems to contradict recently reported deficits of amusic individuals for the processing of verbal material (Patel et al., 2008; Thompson, 2008). However, these deficits have been observed for the processing of intonation and affect in speech. More specifically, the observed difficulties concerned the processing of rapid pitch glides in some amusics (Patel et al., 2008) or the recognition of some emotional expressions (i.e., fear and irritation, Thompson, 2008). In contrast to the processing of this prosodic information, which requires the processing of pitch in the verbal signal, our verbal memory task required memorizing the verbal content of the words and here, amusics' performance was unimpaired.

Amusics' performance on the pitch and word conditions is in agreement with previous studies in showing pitch deficits along with normal performance on verbal memory tests (e.g., digit span). The results on the pitch condition further show that the deficit in congenital amusia is not only linked to a perceptual deficit for pitch, but also to a memory deficit for pitch. The pitch memory deficit has been obtained with tone sequences adjusted as a function of the amusics' pitch thresholds, thus excluding pitch discrimination deficits as the origin of the impaired performance. This finding is in agreement with Foxton et al. (2004) who investigated the perception of a pitch change in a sequence: Increasing the pitch change to exceed pitch threshold did not overcome the amusics' difficulty in detecting differences between two tone sequences. Both data sets thus suggest that while there is a deficit in pitch discrimination at the perceptual level, the amusics' deficit extends to pitch memory and global sequence perception. It might be argued that the impaired pitch performance in the present study might also reflect the amusics' deficit for the perception and recognition of pitch direction (Foxton et al., 2004; Loui, Guenther, Mathys, & Schlaug, 2008), which is not controlled for by pitch differences exceeding discrimination thresholds. However, the amusics' superior performance for contour-violating trials compared to contour-preserving trials suggests some spared pitch direction processing. Specifically, the processing of melodic contour (i.e., the patterns of ups and downs) requires both the processing of pitch direction and pitch memory, which is necessary to link the tones together and enable global sequence perception.

The amusics' impaired pitch memory together with spared verbal memory suggests some dissociation of tonal and verbal information in auditory short-term-memory. This is in agreement with Deutsch's (1970) finding that pitch memory is subject to interference from other pitch information (tones), but not from verbal information (e.g., numbers). While Pechmann and Mohr (1992) reported that this specificity of interference is restricted to musician listeners and does not extend to non-musician listeners, Semal et al. (1996) have specified that pitch memory is influenced more strongly by the proximity of the pitch of the interfering sounds than by the verbal versus non-verbal nature of the interfering material. Their findings led to the hypothesis that there is a specialized memory module for the processing of pitch, which is dissociated from memory for verbal and timbral information (Semal & Demany, 1991; Semal et al., 1996). In contrast to these studies where pitch was the only task-relevant dimension, in our study, verbal and timbre material were also investigated as the task-relevant dimension (i.e., pitch was kept constant for verbal and timbre materials). In the case of verbal material, it is known that short-term memory can benefit from verbal rehearsal strategies (e.g., the phonological loop in the model by Baddeley, 1990), and this process seems to be intact in the amusic individuals. For normal listeners, similar rehearsal processes might apply to short-term memory for pitch (via singing/humming). Some research has suggested that this maintenance of pitch might also benefit from the phonological loop (Williamson, Baddeley, & Hitch, 2006), while others have suggested a specifically dedicated tonal (or musical) loop (Berz, 1995). The amusics' impaired pitch performance might be linked to an impaired tonal loop or to an intact phonological loop failing to apply to impaired pitch memory traces. In contrast to verbal and tonal material, rehearsal processes using internal motor recoding (as in the phonological loop) are unlikely for the short-term memory of timbre material, as the sounds are difficult to reproduce or describe. It would rather appear to be the case that performance on the timbre task reflects sensory-based memory (Crowder, 1993).[3]

The amusics' poor performance for the timbre material suggest that their pitch memory deficit can be accompanied by a memory deficit for other non-verbal material (such as the musical timbres used here). Although the deficit is less marked than for pitch memory and does not affect all amusics, the amusics as a group performed significantly below the control group. In addition, a correlation was observed between the MBEA scores for the contour sub-test and performance on the timbre condition. It is worth noting here that musical timbres contain spectral information that can create a spectral pitch height (Singh & Hirsh, 1992). In sequences, the timbres might create progressions of spectral pitches, from which spectral contours emerge (see also McDermott, Lehr, & Oxenham, 2008). The processing of this information might there-

[3] For our material here, we also suggest that verbalization or labeling is not used for the timbre sequences (even if showing equivalent performance for the word material in control participants). Firstly, the timbres were presented at a relatively fast pace (i.e., 500 ms inter-onset-interval). Secondly, a more demanding task with this material, requiring participants to mentally manipulate the information to find the reverse order, has shown that an additional task during the delay (e.g., counting from 1 to 5) decreases performance for verbal material, but does not alter performance for timbre material (Schulze & Tillmann, 2007).

fore be affected by the pitch deficit. In support of this, the amusics with lower scores on the MBEA contour sub-test also obtained lower scores on the short-term memory test for timbre information. Finally, we need to consider that the amusics' performance on the timbre task might reflect a combination of a timbre memory deficit and a timbre-discrimination deficit because we had not specifically tested for the amusics' perception of the timbres used here. However, previous research suggests no deficit in timbre processing in amusia: Amusic individuals perform normally on tests assessing the recognition of environmental sounds and human voices, both of which require timbre processing (Ayotte et al., 2002; Foxton et al., 2004; Peretz et al., 2002).

It is known that timbre is important for Auditory Scene Analyses (ASA) and the separation of musical 'streams' (Bregman, 1990). When listening to musical pieces, auditory streaming is required in order to separate several melodic lines that are played by the same or different instruments, and both pitch and timbre are known to play a role in this. Foxton et al. (2004) have shown that congenital amusics do not show deficits in ASA based on pitch information. In light of the present findings, ASA performance based on timbre should be tested in congenital amusia. If these listeners have difficulty in keeping track of the melodic lines played by different instruments, the musical structures would become obscured. This would in turn increase the resulting musical deficit, and together with the pitch deficit, would act to hinder the normal acquisition of knowledge about musical structures by mere exposure, as is achieved by normal non-musician listeners (Bigand & Poulin-Charronnat, 2006; Tillmann, Bharucha, & Bigand, 2000).

Acknowledgments

This research was supported by a grant from the Agence Nationale de la Recherche of the French Ministry NT05-3_45978 "*Music and Memory*" to BT. JMF is supported by the Human Frontier Science Program. We thank Olivier Bertrand, Vincent Farget and Géraldine Lebrun-Guillaud for their collaboration on the amusia project in Lyon.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bandc.2009.08.003.

References

Ayotte, J., Peretz, I., & Hyde, K. L. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain, 125*(2), 238–251.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Hove, UK: Lawrence Erlbaum Associates, Ltd..

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception, 12*(3), 353–364.

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.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Conrad, R., & Hull, A. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429–432.

Crowder, R. G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 113–145). Oxford: Oxford University Press.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science, 168*, 1604–1605.

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*, 801–810.

Gosselin, N., Jolicoeur, P., & Peretz, I. (2008). *Impaired memory for pitch*. Paper presented at the Neurosciences and Music III – Conference "disorders and plasticity", Montreal.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15*, 356–360.

Kawahara, H., & Irino, T. (2004). Underlying principles of a high-quality speech manipulation system STRAIGHT and its application to speech segregation. In P. L. Divenyi (Ed.), *Speech separation by humans and machines* (pp. 167–180). Kluwer Academic.

Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In S. Nielzen & O. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43–54). Amsterdam: Excerpta medica.

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology, 18*, R331–R332.

McAdams, S., Winsberg, S., Donnadieu, S., DeSoete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities and latent subject classes. *Psychological Research, 58*, 177–192.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science, 19*, 1263–1271.

Patel, A. D., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition, 59*, 310–313.

Patel, A. D., Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception, 25*, 357–368.

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working memory model. *Memory and Cognition, 20*, 314–320.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., et al. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33*(2), 185–191.

Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key and unaware. *Brain, Advance Access, 31*(March).

Peretz, I., Champod, S., & Hyde, K. L. (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., Gosselin, N., Tillmann, B., Cuddy, L. L., Trimmer, C., Paquette, S., et al. (2008). Online identification of congential amusia. *Music Perception, 25*, 331–343.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Science, 7*(8), 362–367.

Samson, S., Zatorre, R., & Ramsay, J. O. (1997). Multidimensional scaling of synthetic musical timbre: Perception of spectral and temporal characteristics. *Canadian Journal of Experimental Psychology, 51*, 307–315.

Schulze, K., & Tillmann, B. (2007). *Working memory for pitch, timbre and words*. Paper presented at the 15th conference of the European Society for Cognitive Psychology, Marseille.

Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America, 89*, 2404–2410.

Semal, C., Demany, L., Ueda, K., & Hallé, P. A. (1996). Speech versus nonspeech in pitch memory. *Journal of the Acoustical Society of America, 100*, 1132–1140.

Singh, P. G., & Hirsh, I. J. (1992). Influence of spectral locus and F0 changes on the pitch and timbre of complex tones. *Journal of the Acoustical Society of America, 92*, 2650–2661.

Starr, G. E., & Pitt, M. A. (1997). Interference effects in short-term memory for timbre. *Journal of the Acoustical Society of America, 102*, 486–494.

Stewart, L., McDonald, C., Kumar, S., Deutsch, D., & Griffiths, T. D. (2008). *A role for pitch memory in congenital amusia*. Paper presented at the International Conference of Music Perception and Cognition 10, Saporro.

Thompson, W. F. (2008). Impairments of emotional prosody among individuals with amusia. Paper presented at the Neurosciences and Music III Conference, Montreal.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107*(4), 885–913.

Williamson, V. J., Baddeley, A. D. & Hitch, G. J. (2006). Music in working memory? Examining the effect of pitch proximity on the recall performance of nonmusicians. In *Proceedings of the 9th International Conference on Music Perception and Cognition*, 22–26 August, Bologna (pp. 1581–1590).