Brain and Cognition 136 (2019) 103614

Contents lists available at ScienceDirect

Brain and Cognition

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

Short- and long-term memory for pitch and non-pitch contours: Insights from congenital amusia

Jackson E. Graves[a,b,c,*], Agathe Pralus[a], Lesly Fornoni[a], Andrew J. Oxenham[b], Anne Caclin[a,1], Barbara Tillmann[a,1]

[a] Lyon Neuroscience Research Center (CRNL), CNRS, UMR 5292, Inserm U1028, Université Lyon 1, Lyon, France
[b] Department of Psychology, University of Minnesota, Minneapolis, MN, USA
[c] Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale supérieure, PSL University, CNRS, 75005 Paris, France

ARTICLE INFO

Keywords:
Pitch
Brightness
Loudness
Amusia
Melody
Memory

ABSTRACT

Congenital amusia is a neurodevelopmental disorder characterized by deficits in music perception, including discriminating and remembering melodies and melodic contours. As non-amusic listeners can perceive contours in dimensions other than pitch, such as loudness and brightness, our present study investigated whether amusics' pitch contour deficits also extend to these other auditory dimensions. Amusic and control participants performed an identification task for ten familiar melodies and a short-term memory task requiring the discrimination of changes in the contour of novel four-tone melodies. For both tasks, melodic contour was defined by pitch, brightness, or loudness. Amusic participants showed some ability to extract contours in all three dimensions. For familiar melodies, amusic participants showed impairment in all conditions, perhaps reflecting the fact that the long-term memory representations of the familiar melodies were defined in pitch. In the contour discrimination task with novel melodies, amusic participants exhibited less impairment for loudness-based melodies than for pitch- or brightness-based melodies, suggesting some specificity of the deficit for spectral changes, if not for pitch alone. The results suggest pitch and brightness may not be processed by the same mechanisms as loudness, and that short-term memory for loudness contours may be spared to some degree in congenital amusia.

1. Introduction

Pitch, a perceptual correlate of the overall periodicity of a sound, is the auditory dimension that defines melody and harmony in music, influences prosody in speech, and can even alter semantic meaning in tone languages. Pitch contour, or the pattern of changing pitch directions over time, is encoded in short-term memory (STM), allowing for immediate comparison with new stimuli, as well as in long-term memory (LTM), allowing for future recognition and reproduction of familiar melodies. Contour is not the only aspect of melody stored in memory: the interval sizes between successive notes and/or each tone's position within a scale or tonal hierarchy are also encoded (Dowling, 1978), together with the melody's rhythm or timing information. Although some information about the absolute pitches of a melody may persist in LTM (Frieler et al., 2013; Levitin, 1994; Smith & Schmuckler, 2008), melodies are primarily encoded into memory in terms of their relative-pitch attributes, including pitch contour (Plantinga & Trainor, 2005; Schellenberg & Habashi, 2015).

Although melodic contours in Western tonal music are typically defined in terms of pitch, there are some exceptions, such as timbre-based melodies or *Klangfarbenmelodien* (Schoenberg, 1911). Indeed, it has been shown that listeners are capable of recognizing melodic contours in at least two other auditory dimensions, brightness (an aspect of timbre) and loudness (McDermott, Lehr, & Oxenham, 2008). As with pitch, contours in brightness and loudness can be reliably compared within or across dimensions, and can also support the recognition of familiar memories from LTM. They can even produce melodic expectations similar to those found in pitch melodies for both musicians and non-musicians, such as expectations for small intervals and the tendency for melodies to regress towards the mean (Graves, Micheyl, & Oxenham, 2014). All three of these dimensions are known to interact perceptually (Allen & Oxenham, 2014; Marozeau & de Cheveigné, 2007; Melara & Marks, 1990; Verschuure & van Meeteren, 1975; Warrier & Zatorre, 2002), with pitch and brightness perception perhaps relying on overlapping neural mechanisms (Allen, Burton, Mesik, Olman, & Oxenham, 2019; Allen, Burton, Olman, & Oxenham, 2017;

* Corresponding author at: Laboratoire des systèmes perceptifs, École normale supérieure, 29 Rue d'Ulm, 75005 Paris, France.
E-mail address: jackson.graves@ens.fr (J.E. Graves).
[1] Joint last authors.

https://doi.org/10.1016/j.bandc.2019.103614
Received 19 April 2019; Received in revised form 11 September 2019; Accepted 13 September 2019
Available online 20 September 2019
0278-2626/

Bizley, Walker, Silverman, King, & Schnupp, 2009; Warren, Jennings, & Griffiths, 2005). Specifically, changes in pitch and brightness seem to activate strongly overlapping regions of auditory cortex, but produce patterns of activation within those regions that can be statistically distinguished (Allen et al., 2019, 2017). When studied separately, fundamental frequency (F0) and pitch seem most strongly represented in anterolateral primary auditory cortex (Bendor & Wang, 2005; Norman-Haignere, Kanwisher, & McDermott, 2013), whereas spectral shape (resulting in brightness variations) may be more strongly represented in the superior temporal sulcus (Warren et al., 2005). The overlap in neural representations seems unlikely to be due to direct interactions between the physical dimensions; for instance, McDermott et al. (2008) found only small and unreliable direct effects of spectral centroid and intensity on pitch, suggesting that contour perception in brightness and loudness is likely not due to the effect of these dimensions on perceived pitch. The dominance of pitch in defining melodic contours may be because pitch is the structural dimension for tonality and harmony, and because pitch is more accurately coded than either brightness or loudness, with basic difference limens for pitch being an order of magnitude smaller (better) than those for loudness or brightness (McDermott, Keebler, Micheyl, & Oxenham, 2010).

Whatever sets pitch apart from other dimensions in music may also explain why its perception is specifically impaired in congenital amusia. Occurring in about 2–4% of the population (Kalmus & Fry, 1980; Peretz, 2016; Peretz, Cummings, & Dubé, 2007), congenital amusia is a disorder characterized by specific deficits in music perception, despite clinically normal hearing (as defined by pure-tone audiometric thresholds) and normal auditory processing of some non-musical stimuli (Ayotte, Peretz, & Hyde, 2002). One explanation for the specificity of this disorder to music is that it involves deficits in fine-grained pitch perception and pitch memory, whereas rhythm perception (Hyde & Peretz, 2004) and memory for speech stimuli remain intact (Albouy et al., 2018; Tillmann, Schulze, & Foxton, 2009). Nevertheless, the deficits in pitch perception may extend beyond music to the perception of speech prosody (Pralus et al., 2018; Thompson, Marin, & Stewart, 2012) and lexical tones (Jiang, Hamm, Lim, Kirk, & Yang, 2010, 2012; Liu, Jian, Francart, Chan, & Wong, 2017; Liu, Patel, Fourcin, & Stewart, 2010; Nan, Huang, Wang, Liu, & Dong, 2016; Tillmann et al., 2011). The hypothesized deficit of pitch-specific auditory mechanisms makes congenital amusia a topic of interest for pitch researchers, allowing for insights into the perception of consonance (Cousineau, McDermott, & Peretz, 2012) and pitch encoding mechanisms involving frequency components individually resolved by the auditory filters (Cousineau, Oxenham, & Peretz, 2015).

Studies investigating the neural origins of amusia point to abnormalities in connectivity between auditory areas in temporal cortex and later processing in frontal cortex (Albouy, Mattout et al., 2013; Albouy, Mattout, Sanchez, Tillmann, & Caclin, 2015; Hyde et al., 2007; Hyde, Zatorre, Griffiths, Lerch, & Peretz, 2006; Hyde, Zatorre, & Peretz, 2010; Leveque et al., 2016). This frontotemporal network is known to be involved in pitch memory in non-amusic individuals (Zatorre, Evans, & Meyer, 1994). Amusia may also entail a deficit in early encoding of pitch in auditory cortex. Studies using fMRI have suggested that initial cortical encoding of pitch is normal in the amusic population (Hyde et al., 2010; Norman-Haignere et al., 2016), but using the finer temporal resolution of MEG, Albouy, Mattout et al., 2013 observed differences between amusics and controls during early auditory cortical processing. In summary, amusics likely exhibit abnormalities in the initial encoding of pitch as well as abnormalities in frontotemporal connectivity, possibly affecting short-term storage of pitch in working memory.

A number of studies have suggested that the deficits associated with amusia extend beyond the initial encoding of pitch to the representation and retrieval of pitch within STM (for a review, see Tillmann, Lévêque, Fornoni, Albouy, & Caclin, 2016). The observed STM deficit for pitch might also help explain why the pitch discrimination deficit in amusia is especially pronounced in tasks that involve identifying the *direction* of the pitch change, rather than just the presence or absence of pitch change (Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Stewart, 2011; Williamson & Stewart, 2010). This insensitivity to the direction of the pitch change is the basis of the melodic contour deafness hypothesis (Patel, 2010, p. 233), which posits that amusia may involve a specific deficit perceiving pitch contour (largely determined by the successive directions of the pitch changes), which negatively impacts music perception but largely spares speech perception, at least for non-tonal languages.

Amusia may also impact LTM representations of melodies and melodic contours, although it remains unclear whether amusics' difficulty in recognizing familiar melodies (Ayotte et al., 2002) is due solely to poorer initial pitch encoding or a combination of encoding and LTM deficits. Interestingly, although amusics often report themselves to lack an ability to supply the name of familiar well-known melodies, they were able to distinguish unfamiliar from familiar (previously heard and stored in LTM) melodies, as expressed with familiarity judgments on a subjective scale (Tillmann, Albouy, Caclin, & Bigand, 2014).

Finally, recent findings of amusic deficits in tasks such as the detection of amplitude modulation (Whiteford & Oxenham, 2017), have raised some doubts about the purely pitch-specific nature of amusia. Indeed, it has also been suggested that amusia affects the perception and memory of musical timbre (Marin, Gingras, & Stewart, 2012; Tillmann et al., 2009), which might be explained by a deficit in spectral pitch. Nevertheless, in the context of cross-modal integration between vision and audition, some recent evidence (Lu, Sun, Ho, & Thompson, 2017) has suggested that the integration of contour cues (presented in either auditory or visual modality) may be less impaired in amusics for auditory brightness and loudness than it is for pitch. It remains unknown, however, whether amusics' deficits in contour perception are restricted to pitch in the absence of any audiovisual integration task. A better understanding of the extent of the contour perception deficit in amusia is helpful in at least two ways. First, it may help to distinguish which mechanisms are or are not affected in amusia; and second, it may provide further insight into the mechanisms underlying the processing of contours in normal-hearing listeners. A similar strategy has been used in previous studies of pitch perception in amusia (Cousineau et al., 2012, 2015).

The goal of our study was to determine whether amusia-related deficits in melodic contour perception extend to auditory dimensions other than pitch, notably brightness and loudness. To test the pitch specificity of the contour processing deficit associated with amusia, we conducted two experiments, the first primarily involving LTM and the second primarily involving STM. In Experiment 1, participants were asked to identify familiar melodies when represented by changes in pitch, stretched pitch (where the contour was maintained, but the pitch interval sizes were increased), brightness, loudness, or only rhythm. In Experiment 2, participants were asked to compare the contours of novel melodies defined by varying pitch, brightness, or loudness. In both experiments, we evaluated the size of the amusia-related deficit in each of the auditory dimensions in which the stimuli were presented.

2. General methods

2.1. Participants

Twelve amusic individuals (7 female and 5 male) and 12 yoked control participants (7 female and 5 male), matched in terms of age, education, musical experience, gender, and handedness (see Table 1) participated in the study. All participants were screened for normal audiometric hearing thresholds, and congenital amusia was identified with the Montreal Battery of Evaluation of Amusia (MBEA) (Peretz, Champod, & Hyde, 2003). Amusics were initially screened with an online test using the Scale subset of the MBEA, followed with identification using the full MBEA in a laboratory setting. All participants'

Table 1
Demographic information, F0 difference limens (DL) in semitones (ST), and MBEA scores for 24 participants (12 amusics and 12 matched controls). Musical training is defined as the number of years of formal musical education. The cutoff for amusia is defined as two standard deviations below the general population mean of 88% correct, i.e. approximately 78% correct (Peretz et al., 2003). MBEA scores are listed both globally (all six subtests) and for the pitch subtests only (Scale, Interval, and Contour). Mean values are shown, with standard deviations in parentheses. Bold values indicate between-group differences in two-tailed independent *t*-tests with p values < 0.05.

	Amusics	Controls	p value
Age (years)	37.6 (15.6)	37.9 (15.8)	0.96
Education (years)	15.5 (2.8)	15.6 (2.1)	0.93
Musical training (years)	0.08 (0.29)	0.00 (0.00)	0.34
F0DL (semitones)	**1.97 (1.60)**	**0.20 (0.13)**	**0.0028**
MBEA: global (percent)	**71.7 (6.2)**	**89.3 (3.0)**	**< 0.0001**
MBEA: pitch (percent)	**69.0 (7.7)**	**89.4 (5.2)**	**< 0.0001**

pitch discrimination abilities were measured via fundamental frequency difference limens (F0DLs) in the laboratory. All participants provided written informed consent and were compensated for their participation. Study procedures were approved by the required ethics committee, the Comité de Protection des Personnes (CPP), Sud-Est II.

2.2. Stimuli

In both experiments, participants heard melodies made up of harmonic complex tones with Gaussian spectral envelopes, after McDermott et al. (2008). The spectral envelope of each tone was defined by a Gaussian function with the standard deviation equal to 25% of the mean. We also used harmonic complex tones to carry melodies in the loudness condition, as opposed to the noise bursts used in McDermott et al. (2008), in order to minimize potentially distracting differences between the conditions. Each tone was gated on and off with raised-cosine ramps that were 100 ms in Experiment 2, to match the ramp durations in McDermott et al. (2008), but only 10 ms in Experiment 1, so that on and off ramps did not overlap for very short tones in the melodies.

Melodic contour was carried by one of three dimensions. Fig. 1 illustrates the physical dimensions that were manipulated to achieve variations along the desired auditory dimension. When the melodic contour was carried by pitch, the F0 was varied along a range from 150 to 476 Hz; otherwise it stayed constant at 256 Hz. When contour was carried by brightness, the mean of the Gaussian spectral envelope was varied along a range from 315 to 3175 Hz; otherwise it was constant at 1000 Hz. When contour was carried by loudness, the sound level of the complex was varied along a range from 20 to 80 dB SPL; otherwise it was constant at 65 dB SPL. A change of 1 semitone in F0 in the pitch condition corresponded to a change of 2 semitones (or about 12%) in spectral centroid in the brightness condition, and a change in level of 2 dB (Exp. 1) or 3 dB (Exp. 2) in the loudness condition. These values correspond roughly to mean interval discrimination thresholds in each of these dimensions, as measured by McDermott et al. (2010). The smaller step size for loudness in Experiment 1 was due to the wide range of certain melodies used in that experiment, and the relatively narrow range of levels available. All participants completed Experiment 1 first, immediately followed by Experiment 2.

3. Experiment 1: Recognition of contours of familiar melodies

3.1. Rationale

The aim of this experiment was to test the ability of amusic and control participants to recognize familiar melodies, presumably stored in LTM, by presenting them with contours that were created by changes in pitch, brightness, or loudness. All stimuli retained the original rhythm pattern, and performance was compared to a condition in which only rhythm cues were provided. The melody recognition task of McDermott et al. (2008) was modified for this purpose, primarily by making the task easier to avoid chance performance in either the amusic group or the control group (who were all non-musicians with no formal musical training).

Amusics are known to report difficulty recognizing familiar melodies (Ayotte et al., 2002). Nevertheless, they show familiarity judgments of musical material as do controls, suggesting that even amusic individuals have built up a long-term memory representation of a musical lexicon (Tillmann et al., 2014). Amusics' reported explicit recognition difficulty could arise from a deficit in encoding pitch, or it could entail additional deficits in the storage of melodic contours in LTM, independent of the pitch encoding deficit. We expected to observe a deficit in amusics' LTM storage of familiar melodies, even with an easier closed-set recognition paradigm. The goal of Experiment 1 was to test how this deficit would be modulated by presenting melodies in auditory dimensions other than pitch. Such modulations could reflect the combined effect of initial pitch encoding, quality of contour representation in LTM, and ability to use contour cues across dimensions.

3.2. Method

Ten familiar French melodies (i.e., children's songs without words) were selected for this experiment (see Fig. 2, and supplemental audio S1–S18 for examples). The participants were allowed to familiarize

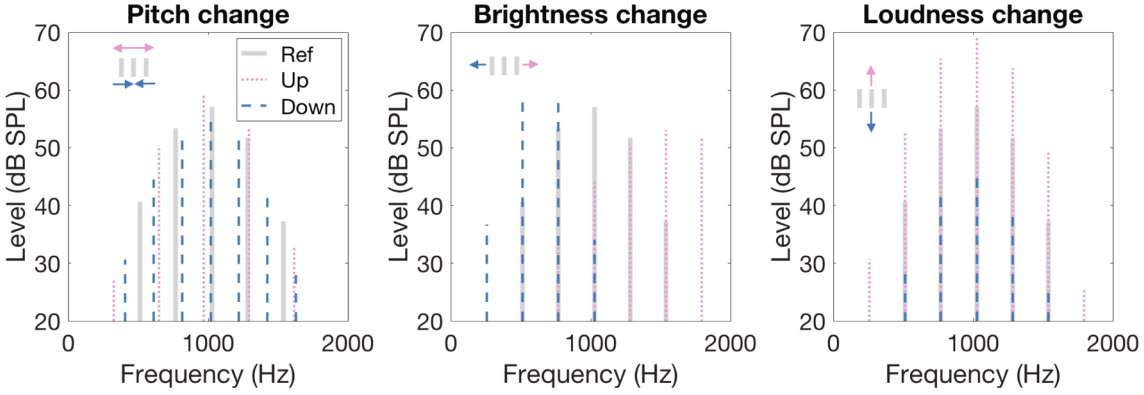

Fig. 1. Spectra of stimuli illustrating shifts in pitch, brightness, and loudness. Left: varying F0 produces a change in pitch. Center: varying spectral centroid produces a change in brightness. Right: varying level produces a change in loudness. In each dimension, the solid lines show the reference tone *Ref* (F0 = 256 Hz, centroid = 1000 Hz, overall level = 65 dB SPL) that was changed upwards (up) or downwards (down): The dotted and dashed lines show increases and decreases, respectively, of 4 semitones (ST) in F0 (left), 8 ST in spectral centroid (center), and 12 dB in level (right).

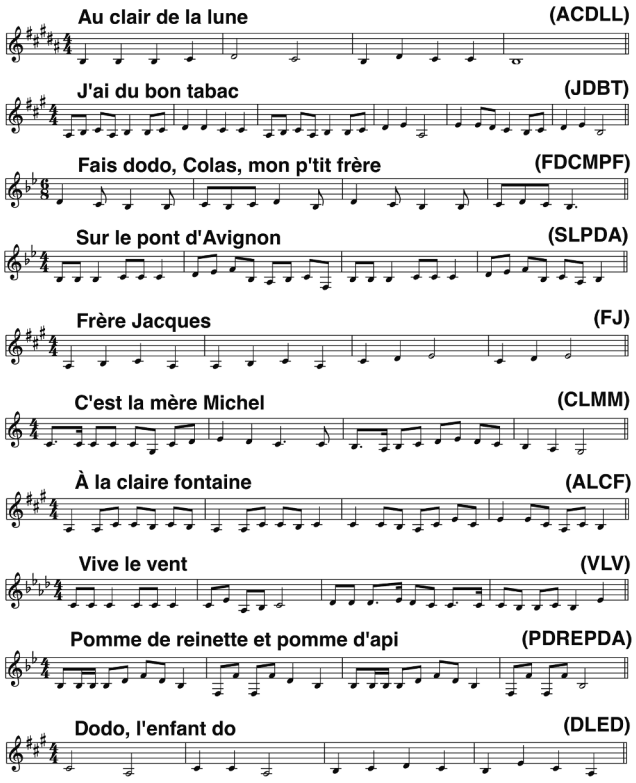

Fig. 2. Musical notation showing the 10 familiar French melodies used for the closed-set recognition task in Experiment 1. All melodies last 9.6 s. The pitch condition is shown here. Participants were instructed to identify these melodies when defined by pitch, stretched pitch (interval sizes doubled), brightness, loudness, or rhythm only. See supplemental audio S1–S18 for examples.

themselves with the melodies before the formal testing, which involved selecting the name of each presented melody from a list of 10 possibilities. Although amusic individuals have been reported to have difficulty recognizing familiar melodies (Ayotte et al., 2002), they also report subjective feelings of familiarity similar to control participants (Tillmann et al., 2014). Furthermore, some aspects of music perception in amusia may be preserved only at an implicit level, sometimes resulting in effects on response time (RT), in the absence of effects on accuracy (Albouy, Schulze et al., 2013). Thus, in addition to our measure of accuracy, we also analyzed two more indirect measures: ratings of confidence, which may reveal differences in levels of awareness, and correct RT, which may shed light on implicit processes that may be spared in amusia.

3.2.1. Stimuli

From a set of melodies previously identified as highly familiar to French university students (Ehrlé, Samson, & Peretz, 2001), and inspired by another study on familiar melody recognition (Devergie, Grimault, Tillmann, & Berthommier, 2010), we chose the ten melodies displayed in Fig. 2. The melodies were generated with the sounds described in the General Methods for the five experimental conditions: pitch (as displayed in Fig. 2), stretched pitch (the same melody, but with all interval sizes doubled), brightness, loudness, or rhythm only. These were a subset of the conditions used by McDermott et al. (2008; Experiment 5). In order to minimize the salience of surface-level cues that would allow participants to perform the task without truly recognizing the melody, the melodies were normalized in terms of tempo and average pitch. Each melody's duration was 9.6 s, usually consisting of 4 measures of 4/4 time at a tempo of 100 bpm. One exception was "*Fais dodo, Colas, mon p'tit frère*", the only melody subdivided in triple time, where the time signature is 6/8, meaning the tempo was 150 bpm for an eighth note or 50 bpm for a dotted quarter note. The other exception was "*J'ai du bon tabac*", which consisted of 6 measures of 4/4 time at a tempo of 150 bpm. In both the pitch and stretched pitch conditions, the absolute pitch of each melody was defined such that the average F0 of the melody was equal to 262 Hz (C4), meaning that the melodies were in different keys from each other, and not tuned to a consistent standard (the notation in Fig. 2 is the closest chromatic approximation). A similar procedure was used for the brightness and loudness conditions, such that the average spectral centroid of all brightness melodies was 1000 Hz, and the average level of all loudness melodies was 65 dB SPL.

3.2.2. Procedure

At the very beginning of the experiment, an example melody that was not used in the actual test ("*Ah, vous dirai-je, maman*") was played in all five conditions, in the following order: pitch, stretched pitch, brightness, loudness, rhythm only. The conditions were not described, but the participant was encouraged to focus on the similarities they heard between the five different versions of the melody, and informed that their task would be to identify melodies played in any of these five styles. After this opening orientation, the names of the ten melodies appeared in a circle on screen, in an adaptation of the paradigm used by Devergie et al. (2010). The participant was asked to listen to each melody in the original "pitch" condition at least once, by clicking on the button next to its name. Then the participant was allowed to replay each melody as many times as desired, in order to ensure familiarity and an ability to connect the name of the song with its melody.

The testing phase, which immediately followed the training phase, consisted of 50 trials. Each of the 10 melodies was presented once in each of the 5 conditions, in a pseudorandom order. The names of the 10 melodies remained on the screen throughout the experiment, but while a melody was playing, the associated response buttons were grayed out. The names were all equidistant from the play button at the center of the screen, where the mouse automatically initialized at the end of each melody. Participants were instructed to respond as quickly as possible once the melody had finished playing, by moving the mouse and clicking on their choice. No feedback was given. After each response, a 5-point Likert scale appeared on which the participants were to rate their level of confidence from 1 ("not at all confident") to 5 ("very confident").

3.2.3. Data analysis

Three dependent variables were recorded on each trial: accuracy (0 or 1), RT (in ms) for correct trials, and confidence (an ordinal rating between 1 and 5). We modeled the variance in each of these 3 dependent variables by fitting a generalized linear mixed model (GLMM), implemented in R using the LME4 package (Bates, Maechler, Bolker, & Walker, 2013). We followed a version of the model selection procedure described by Jaeger (2008), wherein one factor at a time is removed from a large initial model, until the point where removing the factor significantly degrades the model. We followed the rule that a factor was discarded only if the new model without that factor had a lower Bayesian Information Criterion (BIC), a measure of model fit that avoids overfitting by penalizing complexity. The factor chosen to be potentially removed at each step was that with the lowest F-statistic in the current model. The initial model contained four fixed factors: Group (amusic or control), Condition (pitch, stretched pitch, brightness, loudness, or rhythm only), Melody (10 levels, see Fig. 2), and Repetition Number (each melody was repeated 5 times per participant), as well as Participant as a random factor. We also included all potential interactions between Group, Condition, and Block, but not the interactions involving Melody and Repetition number. Ideally the initial model would contain as many interactions as possible; we included the most relevant interactions without rendering the model computationally intractable. For accuracy data, which were transformed with a logit link function, the response distribution was binomial. The distributions

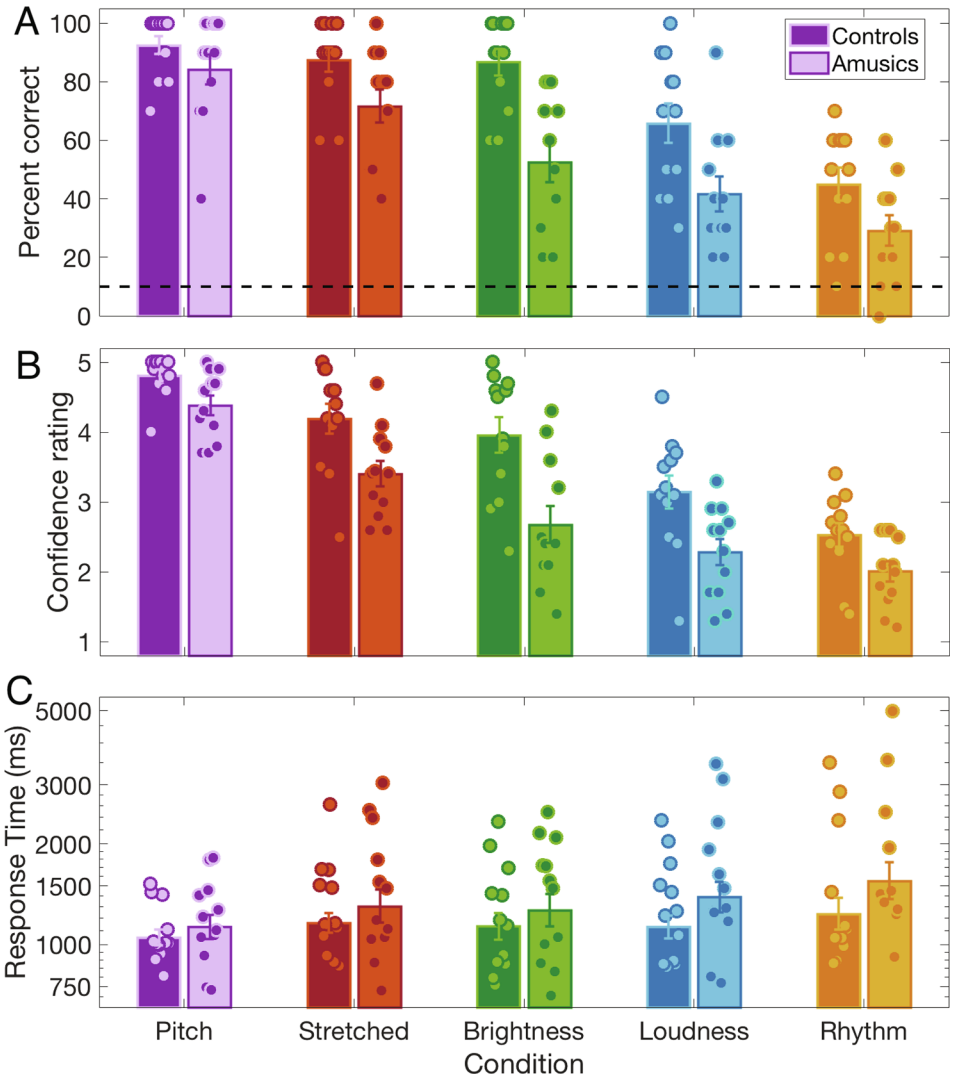

Fig. 3. Results from Experiment 1, by group and condition, for accuracy (A), confidence (B), and RT (C). Throughout, dark bars show means for controls and light bars show means for amusics. Error bars ± 1 SEM; circles represent individual participants. The horizontal dashed line in panel A shows chance performance (10%). RT is plotted only for correct trials, excluding outliers (see Experiment 1, Section 3.2.3).

of the log-transformed RTs and the raw confidence ratings were not significantly different from Gaussian distribution. Analysis of the RTs was limited to correct responses, and outliers were removed before analysis. An outlier was defined as differing from the mean by more than three times the standard deviation of the log-transformed correct response times for each participant. This outlier exclusion process removed between 0 and 2 correct trials per participant (0% to 7.69%, M = 1.75%, S.D. = 2.12%).

3.3. Results

In the familiarization phase, each melody was replayed an average of 1.44 times by amusics (median = 1.45, S.D. = 0.82), and 1.15 times by controls (median = 1.30, S.D. = 0.71). A Wilcoxon signed rank test indicated no significant difference between the median number of replays by amusics and controls (z = 0.81, p = 0.42).

Fig. 3A shows results for accuracy in the main experiment, presented as a function of group and condition. Both groups performed significantly above chance in each of the conditions (one-tailed *t*-tests against chance performance of 10% correct all revealed p-values less than the Bonferroni-corrected α of 0.005 for 10 comparisons).

For accuracy, the final model included fixed effects of Group (Wald χ^2 (1) = 11.75, p = .0006), Condition (Wald χ^2 (4) = 174.51, $p < .0001$), and Melody (Wald χ^2 (9) = 41.05, $p < .0001$), with all other effects and interactions dropped from the model during the model-selection procedure. As defined by Nakagawa and Schielzeth (2013) and calculated in R using the MuMIn package (Bartón, 2018), the proportion of variance explained by the fixed factors, also referred to as marginal R^2, in this final model is 0.317, while the proportion explained by fixed and random factors, also referred to as conditional R^2, is 0.425. The Group effect indicates that amusics were generally impaired relative to the control participants. For Condition, pairwise comparisons revealed that all pairs of conditions, except Pitch vs. Stretched Pitch and Stretched Pitch vs. Brightness, differed significantly from each other after Bonferroni correction (α = 0.005 for 10 comparisons). For Melody, after Bonferroni correction (α = 0.001 for 45 comparisons), *Pomme de reinette et pomme d'api* and *Fais dodo, Colas, mon petit frère* were each recognized more accurately than *À la claire fontaine*, with no other significant comparisons.

Fig. 3B shows results for confidence ratings presented as a function of group and condition. Like the model for accuracy, the final model for confidence included fixed effects of Group (Wald χ^2 (1) = 11.81, p = .0005), Condition (Wald χ^2 (4) = 676.60, $p < .0001$), and Melody (Wald χ^2 (9) = 112.48, $p < .0001$), with all other effects dropped by the model-selection procedure. The variance explained, marginal R^2, by fixed effects in this final model is 0.388, while the variance explained

by fixed and random effects, conditional R^2, is 0.506. The Group effect indicates that amusics were generally less confident in comparison to the control participants. For the factor Condition, pairwise comparisons revealed that every condition differed significantly from every other condition after Bonferroni correction ($\alpha = 0.005$). For the factor Melody, the following comparisons were significant after Bonferroni correction at $\alpha = 0.001$ for 45 comparisons (see Fig. 2 for abbreviations): ALCF < JDBT, FDCMPF, SLPDA, CLMM, VLV, and PDREPDA; ACDLL and DLED < VLV; ACDLL and FJ < PDREPDA.

Fig. 3C shows results for log-transformed RTs in correct trials by group and condition. The final version of the model for log correct RT contained only a fixed effect of Condition (Wald χ^2 (4) = 21.73, $p < .0001$), and no other main effects or interactions. The variance explained, marginal R^2, by fixed effects in this final model is 0.020, while the variance explained by fixed and random effects, conditional R^2, is 0.317. For the Condition factor, pairwise comparisons revealed faster RTs in the Pitch condition than in the Rhythm-only condition, but no other significant differences.

3.4. Discussion

On a closed-set familiar melody recognition task in different auditory dimensions, we observed that amusics were significantly impaired relative to controls, but still performed significantly above chance in all conditions. This deficit is in line with previous findings and supports the claim that difficulty in recognizing melodies without lyrics is a defining feature of amusia (Ayotte et al., 2002). For both groups, performance in the rhythm-only condition was lower than performance in each of the other conditions, confirming that contour information, via pitch, brightness, and loudness, can facilitate melody recognition (McDermott et al., 2008). There was no interaction between group and condition, suggesting that both amusics and controls are capable of using brightness, loudness and pitch cues to extract the contour of a familiar melody, beyond baseline performance based on rhythm cues alone. The lack of an interaction between group and condition suggests that amusics were similarly impaired in their recognition across all dimensions tested. Any pitch-specific deficit in initial encoding, while not directly tested in Experiment 1, could also manifest itself as impairment in all dimensions tested, as all melodies were initially encoded into LTM in the pitch dimension (in everyday life). However, if there is any additional impairment in recognizing contours already stored in LTM, Experiment 1 found no evidence that this impairment might be specific to pitch.

The pitch condition did not produce significantly better accuracy than the stretched-pitch condition, although the mean difference was in the expected direction (with higher mean accuracy for pitch than stretched pitch). This finding does not show a clear benefit of tonal hierarchy cues over the contour-only conditions (i.e., stretched pitch, brightness, and loudness), unlike the results from McDermott et al. (2008), which showed improved performance for the pitch condition over the stretched-pitch condition. This tonal hierarchy effect was possibly not observed in our study because the closed-set task was so easy that both pitch and stretched pitch produced performance close to ceiling.

For RT, no significant effect of group was observed, suggesting that amusics were as quick as controls to respond on correct trials, even though they were less accurate and less confident than controls. In line with previous studies investigating congenital amusia (Albouy, Schulze et al., 2013; Tillmann et al., 2014), the finding that correct RTs of amusics and controls did not differ provides no evidence that amusics experienced any less implicit familiarity with the melodies than control subjects, at least on correct trials. As participants could not respond until the melody finished playing, and responses were generally slow, as they involved pointing and clicking with a mouse, this RT measure may reflect decision times rather than sensory processing times, or it may simply not be sensitive enough to detect a group-level difference. However, it was sensitive enough to detect within-subject differences between conditions. In a future adaptation of the present paradigm, one might allow responding before the end of the melody; in this case, amusic participants might wait for longer, either because they require more information to access LTM or they need more information to feel confident enough to respond (see also Tillmann et al., 2014).

It is possible that some participants may have consciously converted these melodic contours across dimensions, rather than accessing a dimension-general representation of melodic contour. McDermott et al. (2008) also considered this possibility, and asked their participants in a debriefing whether they had verbalized the contour shapes (e.g. “up-down-up-down”). Only a quarter of their participants reported using this strategy, and only on a minority of trials. For our study, we note that RTs were generally below 3 s in both the brightness and loudness dimensions, and that neither of these dimensions significantly differed from pitch in RT. This suggests that participants generally did not take the time to consciously convert the 10 s-long melodic contours from the brightness and loudness dimensions to the pitch dimension. We cannot exclude the possibility that they could have converted smaller excerpts very rapidly, although this seems improbable given the lack of difference in RT between pitch, brightness, and loudness.

Further analyses of the observed main effect of melody revealed that *Pomme de reinette et pomme d’api* produced the highest accuracy, and produced significantly higher confidence and lower RT than the average melody. This is possibly due to its distinctive large interval jumps, including several octave leaps. *Fais dodo, Colas, mon p’tit frère* did not stand out for good or bad performance, despite the fact that it was the only melody in triple, rather than duple, meter. If participants were focusing on meter as a cue, this melody would be very easy to distinguish from the other nine. A previous study found strong effects of rhythmic segregation on familiar melody recognition in the presence of an interleaving distractor melody (Devergie et al., 2010). That study used isochronous versions of familiar melodies in order to remove rhythm cues within the melodies. An interesting future extension of the present study would involve the use of isochronous contour-only melodies (i.e. isochronous melodies in stretched pitch, brightness, or loudness) to evaluate the relative benefits of rhythm cues and contour cues.

Contrary to what one might expect if the deficit involved in amusia were limited to pitch perception, amusics were impaired in all conditions in Experiment 1, even in the rhythm-only condition. This may imply that amusia results in deficits across all dimensions, including rhythm (Foxton, Nandy, & Griffiths, 2006). However, another interpretation is that the deficits in all conditions arose because the original melodies involved LTM representations of pitch. Because musical melodies in everyday life are usually defined by pitch changes, the familiar melodies used in this experiment were likely originally encoded using pitch changes. Therefore, any impairment of pitch processing would lead to a degraded representation of the melody in LTM in the first place, which could then impair recognition of the melody in any auditory dimension. For this reason, Experiment 1 cannot inform us about the extent to which amusics are impaired in brightness or loudness in isolation, because all conditions tested depend on initial encoding in pitch (i.e., of the melodies in LTM). However, the results of Experiment 1 by themselves also do not provide evidence for any pitch-specific LTM deficit in amusia. The observed impairment across all conditions could be explained by an initial pitch-specific encoding deficit that impaired the transfer to LTM, a dimension-general LTM deficit for contours, a dimension-general encoding deficit that limits the degree to which the contours presented in this experiment could be compared with LTM representations, or some combination of these. To more directly compare contour processing of pitch, brightness, and loudness in congenital amusia, Experiment 2 presented novel contours in all three dimensions to test more specifically potential encoding and STM deficits in pitch, brightness and loudness contours.

4. Experiment 2: Contour discrimination for novel melodies

4.1. Rationale

Although the results from Experiment 1 could be explained in terms of a pitch-specific deficit in LTM representations of contour, it did not rule out shorter-term contour encoding deficits that might transcend pitch. The goal of Experiment 2 was to investigate whether amusics are more or less impaired for STM of contours in brightness and in loudness than in pitch. Although amusia is generally understood as a disorder specifically affecting pitch perception (e.g. Cousineau et al., 2015), some recent studies have also found impairment in non-pitch tasks, such as amplitude modulation (Whiteford & Oxenham, 2017) and timbre perception (Marin et al., 2012; Tillmann et al., 2009). It is therefore unclear whether the deficit in pitch contour processing may also extend to brightness and loudness.

To provide a fair comparison, we could not use familiar melodies as in Experiment 1, as these were likely originally encoded by participants in the pitch dimension. Instead, Experiment 2 used new short melodies and modified another task from McDermott et al. (2008, Experiment 1), which involves STM rather than LTM. On each trial, participants were asked to indicate whether the contours of two novel melodies were the same or different. Although McDermott et al. (2008) experiment also involved comparisons between dimensions, we were primarily interested in whether amusia leads to contour perception deficits that are specific to each dimension, so we restricted the conditions in the experiment to within-dimension comparisons, whereby a contour in one dimension is only compared to another contour in the same dimension.

4.2. Method

4.2.1. Stimuli

The stimulus in each trial was a pair of four-note melodies, termed S1 and S2, separated by a 500-ms silent delay. To improve overall performance relative to the task used by McDermott et al. (2008) in their Experiment 1, we reduced the length of the melodies from 5 to 4 tones, and increased the duration of each note from 300 ms to 500 ms (within a melody, notes were presented consecutively without a silent ISI). Increasing the note duration tends to produce better melody recognition in amusic individuals (Albouy, Cousineau, Caclin, Tillmann, & Peretz, 2016). The second melody (S2) was a transposed version of the first (S1), with S2 always in a higher range than S1. For the pitch condition, the highest tone of S1 was 337 Hz, and the highest tone of S2 was 476 Hz. For the brightness condition, the brightest tone of S1 had its centroid at 1588 Hz, and the brightest tone of S2 had its centroid at 3175 Hz. In loudness, the loudest tone of S1 was 62 dB SPL, and the loudest tone of S2 was 80 dB SPL.

4.2.2. Procedure

Each participant completed 9 blocks of 16 trials each, with 3 blocks for each condition (pitch, brightness, and loudness), in pseudorandom order such that no condition was consecutively repeated. This produced 48 total trials per condition per participant. On each trial, participants heard two 4-tone melodies and were asked to judge whether their contours were the same or different, with four response options: “sure different”, “different”, “same”, and “sure same.” Fig. 4 illustrates a “same” and “different” trial.

Each melody contained 4 tones, and thus 3 intervals between the tones, defining the contour. The intervals in S1 (the first melody) were randomly drawn with replacement from the set of [−4 –3 −2 –1 +1 +2 +3 +4] steps. One “step” was defined as 1 semitone in F0, 2 semitones in spectral centroid, or 3 dB SPL in loudness, depending on condition. The intervals in S2 (the second melody) were either identical to those in S1 (same trial) or one of the intervals was different (different trial). For example, on a different trial, S1 may contain the intervals [−2 1 3] and S2 the intervals [−2 1 –3]. The absolute sizes of the intervals were always identical, with only the sign of one of the intervals changing on different trials. In this way, all different trials include a change in contour.

A block contained 16 trials, with 8 same trials (i.e., S2 had the same intervals as S1) and 8 different trials. In the different trial, either the 2nd or 3rd interval was changed in S2 and its initial direction in S1 was either up or down. The changed interval was either a large interval (2, 5, and 6 steps in pitch, brightness and loudness, respectively) or a small interval (1, 4, and 5 steps in pitch, brightness, and loudness, respectively).

4.2.3. Data analyses

In order to account for any potential effects of response bias, the Receiver Operating Characteristic (ROC) was calculated for each participant in each condition. The four potential responses give 3 different criteria: between “sure same” and “same”, between “same” and “different”, or between “different” and “sure different”. For each of these criteria, the hit rate (proportion of “different” trials categorized above the criterion) and false alarm rate (proportion of “different” trials categorized below the criterion) were calculated. The ROC was drawn by connecting these points, and the area under the ROC (AUROC) was calculated as a bias-free measure of sensitivity (Swets, 1973).

4.3. Results

Fig. 5 shows the results of the AUROC analyses by group and condition. One-sample *t*-tests comparing AUROC to the chance level of 0.5 were significant at $p < .05$ for both groups in all conditions, but with Bonferroni correction ($\alpha = 0.008$ for 6 comparisons), amusics were not significantly above chance in the pitch condition ($p = .027$).

A two-way mixed ANOVA on AUROC, with the between-subjects factor of Group (amusic or control) and the within-subjects factor of Condition (pitch, brightness, or loudness), revealed a main effect of Group, $F(1, 22) = 9.63$, $p = .0052$, and a significant Group by Condition interaction, $F(2, 44) = 3.92$, $p = .027$. Post-hoc comparisons with Bonferroni correction ($\alpha = 0.006$ for 9 comparisons) revealed that controls were significantly more accurate than amusics for pitch and brightness, but not for loudness, and that within groups, no two conditions differed significantly from each other (all p values > 0.044). The main effect of Condition was not significant, $F(2, 44) = 0.51$, $p = 0.60$.

4.4. Discussion

Experiment 2 investigated whether the representation of contour in STM was impaired in amusic individuals just for the pitch dimension or also for other dimensions, such as brightness and loudness. The results confirmed the main finding of McDermott et al. (2008) that contour recognition is possible not only for pitch, but also for brightness and loudness. Amusics were also generally able to perform the contour discrimination task at above-chance levels in all dimensions (except when corrected for multiple comparisons: amusics' performance then did not differ from chance level for the pitch dimension). However, a comparison of the group data suggests that amusics may suffer from impaired processing of contour in pitch and brightness, but not in loudness.

The finding of relatively unimpaired loudness contour perception in amusia is notable, given previous findings that amusics are impaired at detecting loudness changes in a sequence of repeating tones (Tillmann et al., 2016). Thus, the normal performance of amusics on the loudness contour discrimination task in the present study may indicate normal STM of loudness information even despite reduced sensitivity to loudness changes overall. This may be due to compensation for the reduced availability of pitch cues for amusic individuals, leading to increased reliance on non-pitch auditory cues. This interpretation would be in agreement with the recent finding that amusics are less impaired at

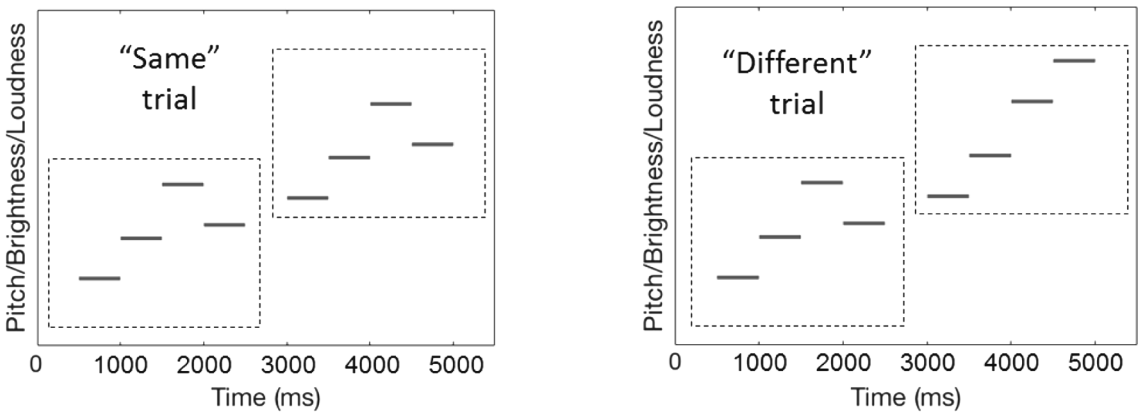

Fig. 4. Schematic illustration of the two trial types for Experiment 2. Half of all trials were "same" trials, where the second melody was an exact transposition of the first. The other half were "different" trials, where one interval was inverted in the second melody. The inverted interval was either the 2nd or 3rd of the three intervals. In this example, the change occurs in the third interval (down in the first melody, up in the second melody).

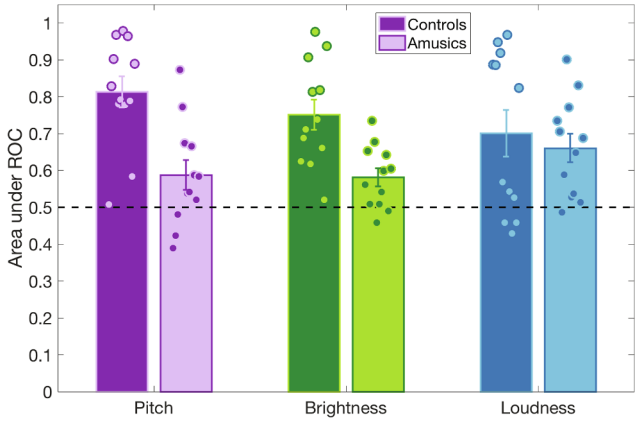

Fig. 5. Area under ROC from experiment 2, by group (dark bars for controls and light bars for amusics) and condition (pitch, brightness, and loudness). Error bars ± 1 SEM; circles represent individual participants. The dashed line indicates chance performance.

audiovisual integration with brightness or loudness than with pitch (Lu et al., 2017). In the present study, however, their lack of impairment was observed within a single dimension, rather than in the context of audiovisual integration.

The observed impairment in amusics for brightness contour perception can be interpreted in two different ways depending on whether brightness is considered to be entirely separate from pitch. Brightness is an aspect of timbre (e.g. McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995), and timbre has been thought of as independent of pitch (ANSI, 1973). Under one interpretation, this finding thus suggests that the deficits in amusia are not specific to pitch: if brightness is separate from pitch, a pitch-specific deficit ought not to impair its perception. Under another interpretation, however, this effect could be seen as further evidence that brightness and pitch are closely related; in fact, pitch and timbre cues are known to interact perceptually (Allen & Oxenham, 2014; Marozeau & de Cheveigné, 2007; Melara & Marks, 1990; Warrier & Zatorre, 2002), and they may activate overlapping brain regions (Allen et al., 2019, 2017; Bizley et al., 2009; Warren et al., 2005). Given these similarities, a deficit in pitch processing should lead to an impairment in brightness processing to the extent that these dimensions are perceived by shared mechanisms. This finding also agrees with previous findings that amusics may exhibit some deficits in timbre processing (Marin et al., 2012; Tillmann et al., 2009). A broad definition of pitch as "spectral pitch", based on changes in tonotopic place of activation, might encompass both pitch and brightness.

There are some timing differences between our study and the previous paradigm on which it is based (McDermott et al., 2008). Notably, in our study the length of the melodies was reduced from 5 to 4 tones, and the duration of each tone was increased from 300 to 500 ms. This was done in order to reduce the overall difficulty of the task, as both manipulations have been shown to improve performance for amusics (Albouy et al., 2016). It is worth noting that, due to the increased tone durations in our paradigm, the relative weights of perceptual encoding and STM maintenance mechanisms on overall performance might have shifted, with slightly decreased weight for perceptual encoding.

5. General discussion

This study compared the ability of amusic and control participants to recognize melodic contours that were presented in the dimensions of pitch, brightness, or loudness. If deficits related to amusia were specific to pitch, then group differences in performance should be heightened when contours are defined by pitch. We observed a constant impairment of amusic participants in all dimensions (i.e., pitch, stretched pitch, brightness, loudness, and rhythm only) when testing long-term recognition (Experiment 1), suggesting that, beyond initial encoding deficits, any additional deficit affecting recognition of contours stored in LTM is likely not specific to pitch. We observed more impairment for short-term retention of pitch and brightness contours than for loudness contours (Experiment 2), suggesting that the deficit at this level may be specific to pitch, or "spectral pitch" if defined broadly to encompass any changes in place of tonotopic excitation, in line with previous findings suggesting timbre deficits in amusia (Marin et al., 2012; Tillmann et al., 2009).

Experiment 1 presented familiar melodies via pitch, brightness, and loudness contours and showed that the performance of the amusic participants was poorer than that of the control group in all conditions. This result could mean that amusics' contour processing of all three dimensions is impaired, but could instead simply be due to the fact that the familiar melodies were originally encoded in LTM mainly via their pitch representations, which were likely impaired in the amusic participants. Experiment 2 tested the participants' ability to distinguish the contours of novel melodies that were presented in terms of either pitch, brightness, or loudness. Here, with melodies processed via STM, an interaction between group and condition was detected, whereby amusics exhibited deficits in the processing of contours presented via pitch and brightness variations, but performed similarly to the control group when the contours were presented via loudness variations. This is despite the fact that performance was above chance for both groups in this condition. Taken together, the results suggest that amusic deficits in contour processing may be specific to pitch, but that the definition of

pitch may extend beyond periodicity (F0) to the percept associated with spectral peaks, known as brightness or spectral pitch (Laguitton, Demany, Semal, & Liégeois-Chauvel, 1998; Oxenham, 2018).

Some previous research has examined deficits in amusia as a way of exploring fundamental mechanisms for pitch perception, on the assumption that the pitch mechanism is specifically impaired in amusia (Cousineau et al., 2012, 2015). If the deficit associated with amusia is indeed linked to fundamental pitch perception mechanisms, our results would suggest a general link between pitch perception and more general spectral processing via the rate-place code (brightness), but not with encoding of level (loudness), even for listeners without amusia. Indeed, the deficit associated with amusia has been previously linked to the rate-place code for pitch: a previous study (Cousineau et al., 2015) reported poorer pitch discrimination thresholds in amusics only for stimuli containing harmonics separately resolved by the auditory filters, which is a requirement for the rate-place code for pitch, but not for the temporal code.

Our study implicates STM for both pitch and brightness as an area of impairment for amusics. The connection between pitch and brightness is in agreement with previous studies finding perceptual interactions between these two auditory dimensions (Allen & Oxenham, 2014; Marozeau & de Cheveigné, 2007; Melara & Marks, 1990; Verschuure & van Meeteren, 1975; Warrier & Zatorre, 2002), and is also consistent with the finding of overlapping regions in auditory cortex that are sensitive to both these dimensions (Allen et al., 2019, 2017; Bizley et al., 2009; Warren et al., 2005). However, the deficit may not necessarily occur at the level of initial encoding, but instead may specifically affect retention of pitch and brightness in short-term memory, through a network connecting frontal and temporal areas (Albouy et al., 2015; Hyde et al., 2010; Leveque et al., 2016).

Despite impaired STM for pitch and brightness, amusics in our study were still able to complete the STM task for loudness in Experiment 2 at a level consistent with controls. This ability may depend on a neural mechanism for encoding intensity, which is separate from those encoding pitch and brightness. For example, a recent study suggests that intensity changes are represented in left Heschl's sulcus (Angenstein & Brechmann, 2013). In order to succeed at this task, amusics may also rely on other contour processing mechanisms, which could be spared in amusia. For example, a recent study showed that manipulation of melodic contours activates the intraparietal sulcus (Foster, Halpern, & Zatorre, 2013), a region associated with visual transformations in space and not typically identified as impaired in amusia.

Due to the nature of STM task used in Experiment 2, where participants were required not only to identify the direction of one pitch change, but to compare the directions of a series of pitch changes, one would predict a large deficit in amusia based on the melodic contour deafness hypothesis (Patel, 2010; Stewart, 2011; Williamson & Stewart, 2010). If this hypothesis is taken to apply to all auditory dimensions, the deficit in amusia should appear in all three of the dimensions tested in Experiment 2. However, if the hypothesis is true only for changes in pitch, and not auditory changes more generally, one might expect to see deficits only in the pitch dimension. Our finding was that pitch and brightness, but not loudness, were impaired on this task. This suggests that it may be specifically the direction of changes in pitch, or a more generally defined spectral pitch that encompasses changes in brightness, that is impaired.

More generally, the results from our study demonstrate that amusic individuals are capable of extracting and storing melodic contours in both STM and LTM. Even if their ability to do so is impaired relative to controls, they share the ability of control participants to extract contour not only from pitch, but also from brightness and loudness patterns. The results of Experiment 1, showing above-chance performance in all conditions, are consistent with those from previous studies (Ayotte et al., 2002; Tillmann et al., 2014), suggesting that amusics may retain some information about familiar melodies in LTM. Together with our results, these findings suggest that even if amusics may have difficulty recognizing familiar melodies without lyrics, they nevertheless retain a degree of implicit memory for these melodic contours that allows them to succeed at this task (as also suggested by the familiarity judgments reported by Tillmann et al., 2014), at least when participants are given a closed set from which to select and/or when indirectly tested with implicit investigation methods.

5.1. Conclusions

Overall, the results of the two experiments suggest that amusia affects processing of contours in pitch and brightness, two closely related dimensions, but not in loudness. These findings suggest that amusia specifically affects pitch contour perception, broadly construed to encompass spectral pitch, or brightness.

Funding

This work was supported by a grant from the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience. This work was supported by LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-0060) and 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-16-IDEX-0005) operated by the French National Research Agency (ANR), and by NIH grant R01 DC005216.

Declaration of Competing Interest

The authors declared that there is no conflict of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bandc.2019.103614.

References

Albouy, P., Cousineau, M., Caclin, A., Tillmann, B., & Peretz, I. (2016). Impaired encoding of rapid pitch information underlies perception and memory deficits in congenital amusia. *Scientific Reports, 6*(18861), 1–11. https://doi.org/10.1038/srep18861.

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., ... Tillmann, B. (2013). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain, 136*(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*, 1–13. https://doi.org/10.3389/fnhum.2015.00020.

Albouy, P., Peretz, I., Bermudez, Pa., Zatorre, R. J., Tillmann, B., & Caclin, A. (2018). Specialized neural dynamics for verbal and tonal memory: fMRI evidence in congenital amusia. *Human Brain Mapping,* 1–13. https://doi.org/10.1002/hbm.24416.

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

Allen, E. J., Burton, P. C., Mesik, J., Olman, C. A., & Oxenham, A. J. (2019). Cortical correlates of attention to auditory features. *Journal of Neuroscience, 39*(17), 3292–3300. https://doi.org/10.1523/JNEUROSCI.0588-18.2019.

Allen, E. J., Burton, P. C., Olman, C. A., & Oxenham, A. J. (2017). Representations of pitch and timbre variation in human auditory cortex. *Journal of Neuroscience, 37*(5), 1284–1293. https://doi.org/10.1523/JNEUROSCI.2336-16.2016.

Allen, E. J., & Oxenham, A. J. (2014). Symmetric interactions and interference between pitch and timbre. *The Journal of the Acoustical Society of America, 135*(3), 1371–1379. https://doi.org/10.1121/1.4863269.

Angenstein, N., & Brechmann, A. (2013). Division of labor between left and right human auditory cortices during the processing of intensity and duration. *NeuroImage, 83*, 1–11. https://doi.org/10.1016/j.neuroimage.2013.06.071.

ANSI (1973). *Psychoacoustical terminology. S3.20.* New York: American National Standards Institute.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain: A Journal of Neurology, 125*(Pt 2), 238–251. http://www.ncbi.nlm.nih.gov/pubmed/11844725.

Bartón, K. (2018). MuMIn: Multi-model inference. R package version 1.42.1.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects

models using Eigen and S4. R package version 1.
Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature, 436*(7054), 1161–1165. https://doi.org/10.1016/0041-0101(79)90119-3.
Bizley, J. K., Walker, K. M. M., Silverman, B. W., King, A. J., & Schnupp, J. W. H. (2009). Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. *The Journal of Neuroscience, 29*(7), 2064–2075. https://doi.org/10.1523/JNEUROSCI.4755-08.2009.
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.
Cousineau, M., Oxenham, A. J., & Peretz, I. (2015). Congenital amusia: A cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia, 66*, 293–301. https://doi.org/10.1002/ana.22528.Toll-like.
Devergie, A., Grimault, N., Tillmann, B., & Berthommier, F. (2010). Effect of rhythmic attention on the segregation of interleaved melodies. *The Journal of the Acoustical Society of America, 128*(1), EL1–EL7. https://doi.org/10.1121/1.3436498.
Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review, 85*(4), 341–354. https://wikis.utexas.edu/download/attachments/3736529/dowling_scale_contour_memory.pdf.
Ehrlé, N., Samson, S., & Peretz, I. (2001). Normes pour un corpus musical. *Annee Psychologique, 101*(4), 593–616. https://doi.org/10.3406/psy.2001.29569.
Foster, N. E. V., Halpern, A. R., & Zatorre, R. J. (2013). Common parietal activation in musical mental transformations across pitch and time. *NeuroImage, 75*, 27–35. https://doi.org/10.1016/j.neuroimage.2013.02.044.
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.
Foxton, J. M., Nandy, R. K., & Griffiths, T. D. (2006). Rhythm deficits in "tone deafness". *Brain and Cognition, 62*(1), 24–29.
Frieler, K., Fischinger, T., Schlemmer, K., Lothwesen, K., Jakubowski, K., & Müllensiefen, D. (2013). Absolute memory for pitch: A comparative replication of Levitin's 1994 study in six European labs. *Musicae Scientiae, 17*(3), 334–349. https://doi.org/10.1177/1029864913493802.
Graves, J. E., Micheyl, C., & Oxenham, A. J. (2014). Expectations for melodic contours transcend pitch. *Journal of Experimental Psychology: Human Perception and Performance, 40*(6), 2338–2347. https://doi.org/10.1037/a0038291.
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, 29*(10), 2562–2570. https://doi.org/10.1093/brain/awl204.
Hyde, K. L., Zatorre, R. J., & Peretz, I. (2010). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex, 21*(2), 292–299. https://doi.org/10.1093/cercor/bhq094.
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. *The Journal of Neuroscience: The Official Journal of the Society for 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.
Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007.Categorical.
Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia, 48*(9), 2630–2639. https://doi.org/10.1016/j.neuropsychologia.2010.05.009.
Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Memory and Cognition, 40*(7), 1109–1121. https://doi.org/10.3758/s13421-012-0208-2.
Kalmus, H., & Fry, D. B. (1980). On tune deafness (dysmelodia): Frequency, development, genetics and musical background. *Annals of Human Genetics, 43*(4), 369–382. https://doi.org/10.1111/j.1469-1809.1980.tb01571.x.
Laguitton, V., Demany, L., Semal, C., & Liégeois-Chauvel, C. (1998). Pitch perception: A difference between right- and left-handed listeners. *Neuropsychologia, 36*(3), 201–207.
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, 116*(1), 88–97. https://doi.org/10.1152/jn.00663.2015.
Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics, 56*(4), 414–423. https://doi.org/10.3758/BF03206733.
Liu, F., Jian, C., Francart, T., Chan, A. H. D., & Wong, P. C. M. (2017). Perceptual learning of pitch direction in congenital amusia: Evidence from Chinese speakers. *Music Perception, 34*(3), 335–351.
Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain, 133*(6), 1682–1693. https://doi.org/10.1093/brain/awq089.
Lu, X., Sun, Y., Ho, H. T., & Thompson, W. F. (2017). Pitch contour impairment in congenital amusia: New insights from the self-paced audio-visual contour task (SACT). *PLoS ONE, 12*(6), 1–15. https://doi.org/10.1371/journal.pone.0179252.
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.
Marozeau, J., & de Cheveigné, A. (2007). The effect of fundamental frequency on the brightness dimension of timbre. *The Journal of the Acoustical Society of America, 121*(1), 383–387. https://doi.org/10.1121/1.2384910.
McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research, 58*(3), 177–192.
McDermott, J. H., Keebler, M. V., Micheyl, C., & Oxenham, A. J. (2010). Musical intervals and relative pitch: Frequency resolution, not interval resolution, is special. *The Journal of the Acoustical Society of America, 128*(4), 1943–1951. https://doi.org/10.1121/1.3478785.
McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science, 19*(12), 1263–1271. https://doi.org/10.1111/j.1467-9280.2008.02235.x.
Melara, R. D., & Marks, L. E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics, 48*(2), 169–178.
Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x.
Nan, Y., Huang, W., Wang, W., Liu, C., & Dong, Q. (2016). Subgroup differences in the lexical tone mismatch negativity (MMN) among Mandarin speakers with congenital amusia. *Biological Psychology, 113*, 59–67. https://doi.org/10.1016/j.biopsycho.2015.11.010.
Norman-Haignere, S. V., Albouy, P., Caclin, A., Mcdermott, J. H., Kanwisher, N. G., & Tillmann, B. (2016). Pitch-responsive cortical regions in congenital amusia. *Journal of Neuroscience, 36*(10), 2986–2994. https://doi.org/10.1523/JNEUROSCI.2705-15.2016.
Norman-Haignere, S., Kanwisher, N., & McDermott, J. H. (2013). Cortical pitch regions in humans respond primarily to resolved harmonics and are located in specific tonotopic regions of anterior auditory cortex. *Journal of Neuroscience, 33*(50), 19451–19469. https://doi.org/10.1523/JNEUROSCI.2880-13.2013.
Oxenham, A. J. (2018). How we hear: The perception and neural coding of sound. *Ann. Rev. of Psychol. 69*(1), 27–50. https://doi.org/10.1146/annurev-psych-122216-011635.
Patel, A. (2010). *Music, language, and the brain.* New York: Oxford University Press.
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., Champod, A. S., & Hyde, K. (2003). Varietes of musical disorders: The Montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences, 999*, 58–75.
Peretz, I., Cummings, S., & Dubé, M.-P. (2007). The genetics of congenital amusia (tone deafness): A family-aggregation study. *The American Journal of Human Genetics, 81*(3), 582–588. https://doi.org/10.1086/521337.
Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition, 98*(1), 1–11. https://doi.org/10.1016/j.cognition.2004.09.008.
Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., & Caclin, A. (2018). Emotional prosody in congenital amusia: impaired and spared processes. *BioRxiv*, 466748.
Schellenberg, E. G., & Habashi, P. (2015). Remembering the melody and timbre, forgetting the key and tempo. *Memory & Cognition, 43*(7), 1021–1031. https://doi.org/10.3758/s13421-015-0519-1.
Schoenberg, A. (1911). *Harmonielehre.* Leipzig and Vienna: Verlagseigentum der Universal-Edition.
Smith, N. A., & Schmuckler, M. A. (2008). Dial A440 for absolute pitch: Absolute pitch memory by non-absolute pitch possessors. *The Journal of the Acoustical Society of America, 123*(4), EL77–EL84. https://doi.org/10.1121/1.2896106.
Stewart, L. (2011). Characterizing congenital amusia. *Quarterly Journal of Experimental Psychology, 64*(4), 625–638. https://doi.org/10.1080/17470218.2011.552730.
Swets, J. A. (1973). The relative operating characteristic in psychology. *Science, 182*(4116), 990–1000.
Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences, 109*(46), 19027–19032. https://doi.org/10.1073/pnas.1210344109.
Tillmann, B., Albouy, P., Caclin, A., & Bigand, E. (2014). Musical familiarity in congenital amusia: Evidence from a gating paradigm. *Cortex, 59*, 84–94. https://doi.org/10.1016/j.cortex.2014.07.012.
Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., & Peretz, I. (2011). Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages. *Frontiers in Psychology, 2*(JUN), 1–15. https://doi.org/10.1007/s00330-015-3615-5.
Tillmann, B., Lévêque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. *Brain Research, 1640*, 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.
Verschuure, J., & van Meeteren, A. A. (1975). The effect of intensity on pitch. *Acustica, 32*, 33–44.
Warren, J. D., Jennings, A. R., & Griffiths, T. D. (2005). Analysis of the spectral envelope of sounds by the human brain. *NeuroImage, 24*, 1052–1057. https://doi.org/10.1016/j.neuroimage.2004.10.031.
Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics, 64*(2), 198–207. http://www.ncbi.nlm.nih.gov/pubmed/12013375.
Whiteford, K. L., & Oxenham, A. J. (2017). Auditory deficits in amusia extend beyond poor pitch perception. *Neuropsychologia, 99*(October 2016), 213–224. https://doi.org/10.1016/j.neuropsychologia.2017.03.018.
Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem. *Memory, 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. *Journal of Neuroscience, 14*(4), 1908–1919.