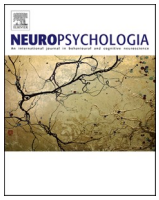
Neuropsychologia 202 (2024) 108960

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia



Deficits in congenital amusia: Pitch, music, speech, and beyond

Caliani Hoarau [a,b,*], Agathe Pralus [a,b], Annie Moulin [a], Nathalie Bedoin [a,c], Jérémie Ginzburg [a], Lesly Fornoni [a], Pierre-Emmanuel Aguera [a], Barbara Tillmann [a,d,1], Anne Caclin [a,1]

[a] Université Claude Bernard Lyon 1, INSERM, CNRS, Centre de Recherche en Neurosciences de Lyon CRNL U1028 UMR5292, F-69500, Bron, France
[b] Humans Matter, Lyon, France
[c] Université Lumière Lyon 2, Lyon, France
[d] Laboratory for Research on Learning and Development, Université de Bourgogne, LEAD-CNRS UMR5022, Dijon, France

ARTICLE INFO

Keywords:
Tone deafness
Pitch perception
Pitch short-term memory
Emotion
Stream segregation
Speech-in-noise
Reading
Writing

ABSTRACT

Congenital amusia is a neurodevelopmental disorder characterized by deficits of music perception and production, which are related to altered pitch processing. The present study used a wide variety of tasks to test potential patterns of processing impairment in individuals with congenital amusia (N = 18) in comparison to matched controls (N = 19), notably classical pitch processing tests (i.e., pitch change detection, pitch direction of change identification, and pitch short-term memory tasks) together with tasks assessing other aspects of pitch-related auditory cognition, such as emotion recognition in speech, sound segregation in tone sequences, and speech-in-noise perception. Additional behavioral measures were also collected, including text reading/copying tests, visual control tasks, and a subjective assessment of hearing abilities. As expected, amusics' performance was impaired for the three pitch-specific tasks compared to controls. This deficit of pitch perception had a self-perceived impact on amusics' quality of hearing. Moreover, participants with amusia were impaired in emotion recognition in vowels compared to controls, but no group difference was observed for emotion recognition in sentences, replicating previous data. Despite pitch processing deficits, participants with amusia did not differ from controls in sound segregation and speech-in-noise perception. Text reading and visual control tests did not reveal any impairments in participants with amusia compared to controls. However, the copying test revealed more numerous eye-movements and a smaller memory span. These results allow us to refine the pattern of pitch processing and memory deficits in congenital amusia, thus contributing further to understand pitch-related auditory cognition. Together with previous reports suggesting a comorbidity between congenital amusia and dyslexia, the findings call for further investigation of language-related abilities in this disorder even in the absence of neurodevelopmental language disorder diagnosis.

1. Introduction

Over the past 20 years, a growing body of research has investigated the phenomenon of congenital amusia, a disorder of music processing. Behavioral studies have revealed a broad range of musical and auditory deficits underpinned by a pitch processing impairment, and neurophysiological studies have revealed an altered (right-sided) fronto-temporal network (for a recent review, see Tillmann et al., 2023). The observed pattern of deficits calls, however, for further investigation, notably regarding speech-in-noise perception, which relies heavily on pitch cues, but has been almost unexplored in congenital amusia, and also regarding language abilities given its sizable comorbidity with dyslexia (Couvignou et al., 2019, 2023; Couvignou and Kolinsky, 2021). The aim of the present study was to provide an overview profile of impairments characterizing congenital amusia, from the pitch-related tasks known to be impaired in congenital amusia to a broader assessment of pitch-related abilities encompassing emotion recognition in speech material, stream segregation with tone sequences and speech-in-noise perception, as well as an assessment of some reading abilities and a subjective assessment of hearing abilities.

* Corresponding author. Centre de Recherche en Neurosciences de Lyon (CRNL, INSERM U1028, CNRS UMR5292, UCBL), équipe Perception, Attention, Mémoire (PAM), 69 675, Bron Cedex, France.
E-mail address: caliani.hoarau@inserm.fr (C. Hoarau).
[1] Equally contributing authors.

https://doi.org/10.1016/j.neuropsychologia.2024.108960
Received 19 January 2024; Received in revised form 17 July 2024; Accepted 17 July 2024
Available online 18 July 2024

1.1. Congenital amusia: a musical disorder linked to a pitch processing deficit

Congenital amusia is a neurodevelopmental disorder characterized by a deficit of music perception and production (Ayotte et al., 2002; Peretz et al., 2002; Peretz and Hyde, 2003). Individuals with congenital amusia have difficulties detecting out-of-key notes in a melody, singing in tune, and recognizing familiar melodies without lyrics. This deficit cannot be explained by hearing loss, brain damage, or lack of music exposure (Peretz, 2013), and is estimated to affect about 1.5% of the population (Peretz and Vuvan, 2017). To detect amusia, Peretz and collaborators developed the Montreal Battery of Evaluation of Amusia (MBEA) (Peretz et al., 2003), providing a common ground for research investigating congenital amusia. In a case study, Lebrun et al. (2012) demonstrated that amusia is also observed in childhood, and a group study revealed that the deficits are similar to those observed in adults diagnosed with congenital amusia (Mignault Goulet et al., 2012).

Over the last two decades, behavioral and neuroimaging studies in congenital amusia have substantially increased our understanding of the behavioral deficits in congenital amusia and their neurophysiological underpinnings (for reviews, see Peretz, 2016; Tillmann et al., 2015, 2023; Williamson and Stewart, 2010). These studies consistently demonstrate that congenital amusia is related to a deficit in pitch perception and memory. Three types of tasks revealed impaired pitch processing in congenital amusia: pitch discrimination/pitch change detection (Albouy et al., 2016; Ayotte et al., 2002; Foxton et al., 2004; Hyde and Peretz, 2004; Peretz et al., 2002), pitch direction/pitch contour processing (Foxton et al., 2004; Loui et al., 2008), and pitch short-term memory (Albouy et al., 2013b; Albouy et al., 2016; Gosselin et al., 2009; Tillmann et al., 2016). For example, in a pitch change detection task, Hyde and Peretz (2004) showed that in contrast to controls who detect changes of a quarter of a semitone in a tone sequence, individuals with amusia have difficulty detecting pitch changes smaller than two semitones. Running adaptive threshold paradigms, individuals with amusia exhibit higher (worse) pitch discrimination thresholds compared to controls even though some participants with amusia have pitch discrimination thresholds that are in the range of controls' thresholds (Foxton et al., 2004; Tillmann et al., 2016). Amusics' impairment has also been shown in a pitch direction task that required participants to determine which of two gliding tones is rising up (Foxton et al., 2004). This impairment in pitch direction identification is also observed with discrete (non-gliding) pitch changes (Loui et al., 2008). Both data sets thus revealed a deficit of pitch contour perception in amusia. Moreover, short-term memory for single tones (Gosselin et al., 2009; Williamson et al., 2010) and for pitch sequences (melodies) is impaired in congenital amusia, impacting each phase of memory processing, from pitch encoding to pitch retention and retrieval (review in Tillmann et al., 2016, 2023). However, verbal short-term memory is preserved in congenital amusia (Albouy et al., 2013a; Tillmann et al., 2009; Williamson et al., 2010). Interestingly, the pitch short-term memory deficit is observed even for individuals with amusia who have a pitch detection threshold similar to controls and even when the pitch change (in the memory task) is larger than the pitch discrimination threshold determined individually for each participant with amusia (Albouy et al., 2013a; Foxton et al., 2004; Tillmann et al., 2009). These data sets confirm that the deficit observed in amusia is not only a pitch perception deficit, but includes a deficit of pitch short-term memory. The pitch memory deficit in participants with amusia is amplified by memory load and interference (Gosselin et al., 2009), as well as by increasing the duration of maintenance (Williamson et al., 2010).

Neuroimaging studies have investigated the neural correlates of congenital amusia with aMRI (anatomical Magnetic Resonance Imaging), fMRI (functional MRI), and EEG/MEG (electroencephalography/magnetoencephalography). Deficits in pitch processing in amusia are associated with anatomical and functional abnormalities in the right inferior frontal gyrus and right superior temporal gyrus (Albouy et al., 2015b; Albouy et al., 2013a; Hyde et al., 2006, 2007; Leveque et al., 2016; Moreau et al., 2013; Peretz et al., 2005). These abnormalities are accompanied by a decreased connectivity between these regions and increased connectivity between the right and left auditory cortices. As a result, the encoding, retention, and retrieval of pitch information are affected (Albouy et al., 2013a; Tillmann et al., 2016, 2023). Specifically, within the right fronto-temporal network, Albouy et al. (2013a) observed reduced and delayed evoked responses during the encoding of melodies, along with abnormal oscillations during the retention delay, and reduced evoked responses during the retrieval of pitch changes.

Overall, congenital amusia is considered a neurodevelopmental disorder with pitch processing deficits related to an impaired fronto-temporal network, resulting in impairments in the music domain given the importance of pitch for melody and harmony. Individuals with amusia exhibit impairments in pitch short-term memory, recognition of musical emotions, singing-in-tune, consonance, tonality, timbre, and sometimes rhythm (for a review see Tillmann et al., 2023). However, pitch processing is also a building block of numerous other auditory abilities, including speech prosody, auditory scene analysis including speech-in-noise perception, and auditory object recognition.

1.2. Beyond music: deficits of pitch-related cognition in amusia

As pitch perception is also relevant for speech processing, some studies investigated how participants with congenital amusia process pitch in speech (e.g., Lu et al., 2015; Patel et al., 2008; Tillmann et al., 2011a, 2011b). Tillmann et al. (2011b) demonstrated that individuals with amusia exhibit deficits when detecting pitch changes in tone or syllable sequences. However, amusics' performance was less impaired for syllable sequences than for tone sequences, revealing slightly smaller deficits when processing pitch in speech sounds compared to musical sounds. The pitch deficit can impact tonal language perception for both non-native and native speakers (Nan et al., 2010; Shao et al., 2020; Tillmann et al., 2011a). Non-tonal-language-speaking amusic participants have difficulty making same/different judgments in a Mandarin Chinese lexical tone discrimination task (Tillmann et al., 2011a). Critically, even mandarin-speaking amusics present impairments in discrimination and identification of mandarin lexical tones in an experimental setting (Nan et al., 2010). The pitch processing deficit in congenital amusia is thus not music-specific, but extends to the processing of pitch in speech. However, these difficulties may have only a limited impact in everyday life of individuals with amusia because pitch changes in speech tend to be larger than in music and individuals with amusia could rely on other acoustic cues (e.g., intensity, duration) for speech comprehension (e.g., Liu et al., 2010; Tillmann et al., 2023).

In speech, pitch also provides relevant information for intentional and emotional prosody (Nooteboom, 1997.; Pihan, 2006; C. Tang et al., 2017). Prosody allows the listener to determine a speaker's intentions and emotions and contributes to non-verbal cue understanding in a conversation. Regarding intentional prosody, individuals with amusia exhibit lower performance compared to controls when tested with statement vs. question discrimination, identification, or imitation tasks based on pitch contour differences in the last word of sentences (Liu et al., 2010). Even emotional prosody processing can be impaired in congenital amusia. Several studies observed a deficit of emotion recognition in speech sounds in amusia (Lima et al., 2016; Pralus et al., 2019; Thompson et al., 2012), in particular for emotion recognition in vowels compared to full sentences (Pralus et al., 2019), suggesting that amusics' difficulties in processing pitch and spectro-temporal parameters affect prosody mostly in cases where the acoustic information is limited and where pitch is the most relevant available cue.

Pitch is essential for other aspects of auditory cognition, such as sound (stream) segregation (Oxenham, 2008), which is the ability to group together spectral energy produced by one sound source and separate it from energy produced by other sound sources. This stream

segregation process notably allows us to understand speech in noise (Hong and Turner, 2006; Oxenham, 2008). This leads to the hypothesis that deficits in pitch perception could cause a streaming deficit in congenital amusia and affect hearing-in-noise abilities. To the best of our knowledge, only one study has investigated sound segregation in congenital amusia. Foxton et al. (2004) used a stream segregation task that assesses the assignment of consecutive sounds alternating in pitch to one single or two distinct perceptual sources (streams) and found no difference between amusic and control participants, suggesting preserved stream segregation abilities in amusia. Subsequently, investigating hearing-in-noise abilities in congenital amusia, studies found divergent results revealing either no deficits of tone or vowel perception in noise in amusics (Loutrari et al., 2024; W. Tang et al., 2018), or reduced speech intelligibility in noise (Liu et al., 2015a; Shao et al., 2016). However, the studies that reported significant differences between amusic and control participants (Liu et al., 2015a; Shao et al., 2016) used lexical tones which rely on fine-grained pitch processing for their identification or discrimination, and deficits in lexical tone processing were already observed in quiet in the same participants. Looking more closely into inter-individual differences, W. Tang et al. (2018) observed a deficit in speech-in-noise processing for lexical tones in a subgroup of congenital amusic participants that are also impaired in lexical tone processing (tone agnosics).

Overall, several data sets suggest that pitch processing in speech sounds is altered in congenital amusia, impacting in particular lexical tone processing and prosody processing. However, there is no agreement yet whether auditory scene analysis, including speech-in-noise perception, is affected in congenital amusia, despite the established importance of pitch cues in such tasks.

1.3. Beyond amusia: relationships with other neurodevelopmental disorders

Comparing congenital amusia to other neurodevelopmental disorders associated with auditory cognition deficits could yield valuable insights in the understanding of neurodevelopmental disorders and more generally brain functioning. Recent studies have revealed some comorbidity between amusia and dyslexia (Couvignou et al., 2019, 2023; Couvignou and Kolinsky, 2021). Couvignou et al. (2019) demonstrated that self-reported dyslexic adults have lower melodic skills than have controls, suggesting a possible connection between these two disorders. Via a series of tests involving reading, verbal working memory, phonological awareness, and musical abilities, they reported that 30% of adult dyslexic participants show a profile of congenital amusia, while 25% of adult congenital amusic participants experience reading difficulties. In children, 34% of dyslexic children present pitch perception and production impairments (Couvignou and Kolinsky, 2021). Children with comorbid dyslexia and congenital amusia have reduced memory skills for both verbal and musical materials compared to typically-developing children and children with dyslexia without amusia, notably in tasks requesting participants to process serial order (Couvignou et al., 2023). Both congenital amusia and dyslexia have a genetic component and are associated with fronto-temporal abnormalities, for which abnormal neural migration has been discussed for both conditions (Boets et al., 2013; Hyde et al., 2006; Peretz, 2016). These findings suggest that memory deficits, in particular those including serial order processes, might define a potential link between the disorders. They call for a more systematic assessment of language abilities (in particular reading) in congenital amusia, and conversely of musical abilities in dyslexia and developmental language disorders. Additionally, links between neurodevelopmental disorders might be found even across sensory modalities: when investigating the comorbidity between congenital amusia and developmental prosopagnosia, Corrow et al. (2019) observed impaired pitch perception deficits in three out of twelve subjects with prosopagnosia.

1.4. The present study

The first aim of the present study was to characterize more precisely the pattern of pitch-related deficits in congenital amusia, combining perception and memory tests classically used in amusia research with tests assessing other aspects of pitch-related auditory cognition in the same participants. Pitch-related auditory abilities are most often studied in isolation (in congenital amusia studies, but also more generally in research investigating auditory processing), our study thus aims to not only shed light on the pattern of deficits in congenital amusia, but also on the organization of auditory perception and cognition more generally. We further explored text reading/copying abilities, which have received less attention up to now in congenital amusia, but which seem relevant to study given the recently documented comorbidity with dyslexia (Couvignou et al., 2019; Couvignou and Kolinsky, 2021). This last point was expected to bring not only new knowledge on congenital amusia per se but also on the interrelations between music and language abilities.

The following set of behavioral tests were used. To target pitch processing in non-verbal material, we measured pitch discrimination thresholds (PDT, Tillmann et al., 2009), pitch change detection (Albouy et al., 2015a; Hyde and Peretz, 2004; Pralus et al., 2021), pitch direction of change identification (Foxton et al., 2004; Loui et al., 2008; Pralus et al., 2021), and short-term memory of pitch sequences (Pralus et al., 2021; Tillmann et al., 2016). To test pitch processing with verbal material, we measured emotional prosody perception (Pralus et al., 2019). Beyond these more classical tasks, we tested stream segregation with tone material (Foxton et al., 2004; Pralus et al., 2021) and verbal material using a speech-in-noise task (Moulin et al., 2013). This assessment was completed with questionnaires related to hearing abilities (Moulin et al., 2015, 2019), and listening effort (Ferschneider and Moulin, 2023). In addition, participants performed text reading and copying tests that are sensitive to variability even among expert-reading adults (Duplat and Girier, 2006; Vialatte et al., 2023). And, finally, three visual control tests measured the ability to switch between local and global attention (Bedoin, 2017; Bedoin and Medina, 2013).

Based on previous findings (reviewed in Tillmann et al., 2016), participants with congenital amusia (identified by their low MBEA scores) were expected to be impaired in pitch processing and memory, leading to low performance scores not only in the PDT, but also in the pitch change detection, pitch direction of change identification, and pitch short-term memory tasks. In these four tasks, participants need to process the pitch of each tone and keep the tones in memory to compare them, yet with different time constraints and instructions focusing on different aspects of pitch (change, contour). For the other pitch-related tasks, we expected to observe deficits in emotional prosody recognition in vowels, but not in sentences (Pralus et al., 2019). As pitch perception is essential in stream segregation including speech-in-noise perception, individuals with amusia could exhibit deficits in the two tasks assessing these abilities with either tone or verbal material and/or differences in subjective reports (questionnaires) compared to control participants. However, the currently available experimental evidence is limited and not congruent (Foxton et al., 2004; Liu et al., 2015a; Loutrari et al., 2024; W. Tang et al., 2018). Furthermore, poorer reading and copying performance could be expected at least in some individuals with amusia because of the comorbidity reported with dyslexia (note,

however, that we included in the present study a homogeneous group of participants not reporting any previously diagnosed dyslexia or developmental language disorder). Finally, for the visual control tasks, no deficits were expected as no visuo-spatial attention deficits have been reported previously in congenital amusia, and spatial processing deficits initially reported (Douglas and Bilkey, 2007) were not further confirmed (Tillmann et al., 2010; Williamson et al., 2011).

2. General methods

2.1. Participants

Prior to the main testing session, participants underwent screening tests including standard audiometry from 250 to 8000 Hz, the MBEA (Peretz et al., 2003), and a measure of their PDT with a two-alternative forced-choice task (Tillmann et al., 2009). These screening tests allowed us to identify participants as amusics or controls (see below), and to exclude participants with hearing loss (thresholds above 20 dB HL at frequencies at or below 4000 Hz). The target number of participants was chosen a priori on the basis of previous lab-based studies of congenital amusia, which typically include between 10 and 20 participants per group (e.g., Hyde et al., 2007; Albouy et al., 2016; Pralus et al., 2019; Tillmann et al., 2011a,b). Individuals with amusia are difficult to recruit given its low prevalence and lack of knowledge about the disorder in the general population, and for the present study we managed to include 18 amusic participants and 19 controls (see Table 1 for demographic information and Table 2 for comparison with sample sizes of previous studies).[2] The two groups were matched for age, sex, education, and musical education. Two participants with amusia and one control participant presented at least one audiometric threshold greater than 35 dB at 6000 or 8000 Hz, which is in line with the age range covered (up to 75 years-old). Participants self-reported normal or corrected to normal vision. None of them reported neurological or psychiatric diseases, nor having a diagnosis of learning disorder such as dyslexia.

Groups were defined as follows: The MBEA is composed of six subtests that assess different components of music perception and memory (scale, contour, interval, rhythm, meter, and memory). To be considered as amusics, participants had to obtain a global score inferior to 23 (maximum score = 30) and/or a MBEA pitch score (average of the first three subtests) inferior to 22 (maximum score = 30). We included in the amusic group one participant with a borderline MBEA global score of 23.6 and a MBEA pitch score of 23 (this participant had an elevated PDT of 1.32 semitones, note that the highest PDT in controls was 0.84 semitone). Data from one additional participant with a MBEA global score of 24.6 and a MBEA pitch score of 23.8 was excluded as these scores were superior to the usually accepted thresholds to consider a participant as amusic, even though inferior to the mean minus two standard deviations of the control group.

The study procedures were approved by an ethics committee (CPP Ile de France VI, ID RCD 2018-A02670-55). Participants gave their written informed consent and were paid for their participation.

Table 1
Demographic data and performance on basic screening tests for amusic and control participants (mean ± standard deviations together with minimum (Min) and maximum (Max) scores within each group were reported). The groups were compared with Student *t* tests (two-sided), except for the sex-ratio n and handedness comparison where a Chi2 test was used. F: female; M: male. R: right-handed; L: left-handed. For the audiometry, we calculated for each participant the mean audiometric thresholds from 250 to 4000 Hz and from 6000 to 8000 Hz across both ears (according to the recommendation of the International Bureau for Audiophonology). MBEA: Montreal Battery for the Evaluation of Amusia (Peretz et al., 2003). PDT: Pitch Discrimination Threshold (Tillmann et al., 2009).

Group	Amusics (n = 18)	Controls (n = 19)	p-value (group comparison)
Age (years)	31.94 (±11.86) Min: 18 Max: 58	32.47 (±13.03) Min: 19 Max: 75	0.90
Sex	12 F, 6 M	13 F, 6 M	0.91
Education (years)	15.11 (±1.71) Min: 12 Max: 17	15.84 (±1.64) Min: 13 Max: 20	0.19
Musical education (years)	0	0	
Handedness	15R, 3L	16R, 3L	0.94
Audiometric thresholds (dB):			
250 to 4000 Hz	8.47 (±3.80) Min: 0 Max: 18.5	7.92 (±5.40) Min: 0.5 Max: 18	0.39
6000 to 8000 Hz	14.65 (±12.16) Min: 2.50 Max: 46.25	12.22 (±5.78) Min: 3.75 Max: 22.5	0.72
MBEA global score (maximum score = 30)	21.66 (±2.07) Min: 17.30 Max: 24.31	26.75 (±1.42) Min: 24.30 Max: 29.50	<0.001
MBEA pitch score (maximum score = 30)	20.25 (±2.14) Min: 16.30 Max: 23.00	27.13 (±1.28) Min: 25.00 Max: 29.60	<0.001
PDT (semitones)	1.10 (±0.80) Min: 0.20 Max: 2.93	0.35 (±0.23) Min: 0.01 Max: 0.84	<0.01

2.2. Behavioral tests

On a separate day from the screening session, participants performed the following tests during the same testing session, which lasted about 90 min in total (see Procedure for task order).

2.2.1. Auditory tests

All auditory tests were run on an iPad touch tablet with two loudspeakers (Logitech Z200). Participants performed six auditory tasks: Pitch Change Detection (PCD), pitch Direction of Change Identification (DCI), pitch Short-Term Memory with four-tone sequences (STM), Auditory Stream segregation (AS), Emotion recognition in full sentences (EMO), and speech-in-noise perception (Audimots). In addition, among the 38 participants, 14 amusic participants and 6 control participants also performed a pitch Short-Term Memory task with six-tone sequences (STM_6, presented after STM) and an Emotion recognition task with vowels (EMO_v, presented after EMO). See Specific Material and Results for details about each task.

The tests PCD, DCI, STM, AS, and EMO were taken from Pralus et al. (2021) (but with all trials being presented audio-only). EMO and EMO_v were used by Pralus et al. (2019). STM_6 was designed for the present study to be matched as closely as possible to the STM task with four-tone sequences.

For PCD, DCI, STM, and STM_6, the stimuli were synthetic tones with twelve harmonics, each lasting 500ms and presented with an Inter-Stimulus-Interval (ISI) of 100ms. The same harmonic tones were used for AS but with a duration of 100ms and an ISI of 20ms (Pralus et al., 2021). For EMO, the stimuli were semantically neutral sentences pronounced by male or female talkers (Pralus et al., 2019, 2021). For EMO_v, /a/ vowels were used, all pronounced by women (Pralus et al., 2019). For Audimots (Moulin et al., 2013), the stimuli were monosyllabic words pronounced by women or men.

For all auditory tests, participants had unlimited time to answer.

[2] After the testing session described here, the participants were then enrolled in a longer-scale training study (with auditory and visual training). The effects of training will be described in a subsequent manuscript.

Table 2
Tasks and reported statistics from previous studies in congenital amusia investigating the three pitch tasks (PCD, DCI, STM), ASA or speech-in-noise measures, and emotional prosody. Cohen's d or Cohen's f were recalculated from the statistical results presented in the studies. It should be noted that some of the cited studies are based on overlapping participant pools. Direct comparisons between studies remain tentative as most of the tasks were different in details of their implementation from the tasks and materials used here.

	N (amusic, control)	Task used	Reported statistic	Recalculated Cohen's f (for ANOVAs) or d (for t-tests or Mann-Whitney-Wilcoxon tests)
PCD (5-tone sequences, 4th tone same or different)				
Hyde and Peretz (2004)	10, 10	1/2, 1, 2, 4, 6 semitones changes	$F(4,72) = 95.4$, $p < 0.0001$ [group x pitch distance interaction]	Cohen's f = 2.21
Albouy et al. (2015)	16, 16	1/8, 1/4, 1/2 semitone changes	$F(1,29) = 4.53$; $p = 0.04$; $MSE = 2.78$; $\eta_p^2 = 0.13$ [group main effect]	Cohen's f = 0.33
DCI				
Loui et al. (2008)	6, 6	2 tones, 2nd lower or higher, 1st at 500Hz and 2nd ranged from 450 to 550Hz	$t(10) = 5.6$, $p = 0.0002$	Cohen's d = 3.54
Liu et al. (2012)	19, 19	Discrete/gliding pitches, rising/falling, 0.01 to 10 semitones	gliding complex tone (controls: 0.19 (0.08), amusics: 1.90 (2.99), $W = 61$, $p = 0.0005$), and discrete complex tone (controls: 0.30 (0.40), amusics: 4.44 (3.18), $W = 14$, $p < 0.0001$)	Cohen's d = 0.81 Cohen's d = 1.83
Liu et al. (2017)	20, 20	2 tones, down to up, up to down, 0.01 to 12 semitones	$t(28) = 6.02$, $p < 0.001$, $d = 2.33$	Cohen's d = 2.33
MCT (pitch sequences, same-different task)				
Ayotte et al. (2002)	10, 20	6 to 15 items per sequence, 1 semitone differences for different pairs	$F(1,28) = 162.89$, $p < 0.001$ [group main effect]	Cohen's f = 2.32
Foxton et al. (2004)	10, 10	4 item per sequence, 2 tones differences for different pairs	$F(1,18) = 75.3$, $p < 0.01$ [group main effect]	Cohen's f = 0.81
Gosselin et al. (2009)	9, 9	3 or 5 items per sequence, permutation of two tones for different pairs	$F(1, 12) = 11.43$; $p < 0.01$ [group main effect]	Cohen's f = 0.86
Tillmann et al. (2009)	10, 10	6 items per sequence, permutation of two tones for different pairs	$F(1,18) = 27.32$, $MSE = 0.05$, $p < 0.0001$ [group main effect] $F(2,36) = 9.06$, $MSE = 0.04$, $p = 0.0007$ [group × material interaction] (Material = music or verbal)	Cohen's f = 1.15 Cohen's f = 0.64
Williamson and Stewart (2010)	14, 14	2 to 5 items per sequences, permutation of two tones for different pairs	$F(1, 26) = 4.62$, $MSE = 8.49$, $p = 0.04$, $\eta_p^2 = 0.15$ [interaction between group and task] (Task = tone or digit span)	Cohen's f = 0.37
Albouy et al. (2013a)	9, 9	6 items per sequence, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 semitones differences for different pairs	$F(1,16) = 75.37$, $p < 0.0001$ [group main effect]	Cohen's f = 2.03
Albouy et al. (2013b)	11, 11	5 items per sequence, permutation of two tones for different pairs	$F(1,20) = 37.38$, $p < 0.0001$ [group main effect]	Cohen's f = 1.28
Albouy et al. (2016)	10, 10	3 or 4 items per sequence, >3 semitone differences for different pairs	$F(1,18) = 6.02$; $p = 0.02$; $MSE = 1838.0$; $\eta_p^2 = 0.25$ [group main effect] [[$F(1,18) = 5.49$, $p = 0.03$, $MSE = 192.8$; $\eta_p^2 = 0.23$ [group × tone duration interaction] (Tone duration: 100 or 350 ms)	Cohen's f = 0.50 Cohen's f = 0.47
Albouy et al. (2019)	18, 18	3 items per sequence, 3, 4, 5, 6, 7, 8 semitones differences for different pairs Comparison with verbal sequences and with a perceptual control task	$F[1,34] = 7.25$; $p = 0.01$ [group × material × task interaction] (Material = music or verbal; Task = memory or perception)	Cohen's f = 0.42
Graves et al. (2019)	12, 12	4 items per sequence, transposition for different pairs	$F(1,22) = 9.63$, $p = 0.0052$ [group main effect] $F(2,44) = 3.92$, $p = 0.027$ [group × condition interaction] (Condition = pitch, brightness, or loudness)	Cohen's f = 0.6 Cohen's f = 0.35
Auditory Scene Analysis				
Foxton et al. (2004)	10, 10	Series of alternating high and low pure tones (triplets)	$F(1,17) = 0.05$, $p > 0.05$ [group main effect]	Cohen's f = 0.05
Speech in noise				
Liu et al. (2015)	16, 16	Stimuli: Sentences (Mandarin); SNR: no noise, -5 dB, 0, +5 dB	$F(1,30) = 15.96$, $p < 0.001$, $\eta_p^2 = 0.16$ [group main effect]	Cohen's f = 0.68
Shao et al. (2016)	14, 14	Stimuli: Words (Cantonese); SNR: no noise, 0, −10 dB	Identification task: $F(1, 26) = 19.050$, $p < 0.001$ [group main effect] Discrimination task: $F(1, 26) = 8.954$, $p = 0.006$ [group main effect]	Cohen's f = 0.8 Cohen's f = 0.53
W. Tang et al. (2018)	12, 12	Stimuli: Vowels (Mandarin), tones; SNR: 4 dB, −8 dB	Quiet: $F(1,29) = 2.018$, $p = 0.166$ [group main effect] Noise: $F(1,29) = 1.613$, $p = 0.214$ [group main effect]	Cohen's f = 0.18 Cohen's f = 0.14
Loutrari et al., 2024	27, 27	Stimuli: Sentences (English), SNR: 5 dB or in quiet	$F(1, 50.91) = 1.61$, $p = 0.21$, $\eta_p^2 = 0.03$ [group main effect]	Cohen's f = 0.11
Emotional prosody (emotion categorization task)				
Thompson et al. (2012)	12, 12	Full sentences	$F(1, 22) = 9.11$, $p = 0.006$, $\eta_p^2 = 0.29$	Cohen's f = 0.58
Lolli et al. (2015)	9, 31	Full sentences	$t(38) = 0.58$, n.s	Cohen's d = 0.19
Pralus et al. (2019)	18, 18	Full sentences	$F(1,34) = 2.24$, $p = 0.14$	Cohen's f = 0.18
Pralus et al. (2019)	18, 18	Vowels	$F(1,34) = 12.5$, $p = 0.001$	Cohen's f = 0.56

Cohen's d: 0.2 = small, 0.5 = medium, 0.8 = large.
Cohen's f: 0.1 = small, 0.25 = medium, 0.4 = large (Cohen, 1992).

2.2.2. Reading and copying tests

The testing battery also included a text reading and a copying test. For the reading test (DeltaText, Vialatte et al., 2023), participants had to read a text aloud as rapidly and correctly as possible. For the copying test, "la Baleine Paresseuse" (meaning "the Lazy Whale", see (Duplat and Girier, 2006), participants had to copy a text hung on the wall as correctly as possible in 3 min.

2.2.3. Visuo-spatial tests

We used visuo-spatial attention tests implemented on an iPad touch tablet as control tasks. Participants performed seven visuo-spatial tasks: five tasks from the Switchipido battery: Switchipido arrow global, Switchipido arrow alternance, Switchipido triplet global, Switchipido triplet local, Switchipido triplet complex, and two tasks from SIGL: local and global (see Bedoin, 2017; Bedoin and Medina, 2013 for details). Instructions were given before each task. For all tests, participants were instructed to respond as fast as possible by clicking on the tablet screen. Participants had unlimited time to give their answer and the stimuli and response buttons remained on the screen until the participant answered.

2.3. Procedure

The experiment took place in a quiet room. As mentioned above, screening tests (audiometry, MBEA, PDT) were performed on a separate day before the main testing session. During the testing session, the participant was seated in front of an iPad touch tablet with two loudspeakers (Logitech Z200) at 70 cm from the participant's head, with 40 cm between the two speakers. The volume was adjusted to be comfortable, except for Audimots where the volume was fixed (see below). Before each test, the participants received an oral explanation of the corresponding task and performed a short training in order to ensure that they understood the task. They received no feedback during the tests. Tests were done in the same order for all participants: auditory tests (PCD, DCI, STM, STM_6, ASA, EMO, EMO_v, Audimots) on the tablet, visuo-spatial tests (Switchipido tests: arrow-global, arrow-alternance, triplet-global, triplet-simple, triplet-complex; SIGL-local, SIGL-global) on the tablet, DeltaText reading, "Lazy Whale" text copy, hearing quality questionnaires on paper. Because our aim was to validate the auditory tests as a listening battery, the same random order of trials was used for all participants.

2.4. Data analyses

Statistical analyses were conducted with the software JASP (JASP 0.14.1). They were first performed separately for each task, with repeated measure analysis of variance (ANOVA) with Group (amusics, controls) as a between-participants factor, Student t-tests or Mann-Whitney tests. Task-specific measures and within-participant factor(s) for ANOVAs will be detailed in the Results section for each of the tests. Note that for auditory tests, response times were calculated for the correct responses from the end of the stimulus. Visual inspection of the data did not reveal very long or too early response times, and no response times were removed. We applied Greenhouse-Geisser correction when the assumption of sphericity was violated. For significant effects and interactions, post-hoc comparisons were calculated with t-tests with Holm correction for multiple comparisons. For group comparisons on single variables, we used Mann-Whitney tests when data were not normally distributed for at least one group. For auditory, reading, and copying tests, one-sided tests were used testing the hypothesis of poorer performance in the amusic group. In a second step, results of all behavioral tests and questionnaires were analyzed jointly with a Principal Component Analysis using the factorMineR package (Lê et al., 2008) in R Studio. We analyzed the loadings of each variable to the components retrieved by the PCA to qualitatively describe the links between the different variables. Finally, we compared between groups the coordinates of the participants on the PCA components with Student *t* tests corrected with Bonferroni post-hoc tests to further analyze the pattern of between group differences.

Power analyses. We inspected the effect sizes of previous studies in congenital amusia using one (or more) of the three pitch tasks (PCD, DCI, STM), speech-in-noise or ASA measures, and emotional prosody (see Table 2). For pitch tasks and emotion recognition in vowels, medium to large effect sizes for between-group differences were reported in previous studies (with mostly large effect sizes), whereas for auditory scene analysis and speech-in-noise results are variable across studies: negligible to small group effects were reported in three studies (Foxton et al., 2004; Loutrari et al., 2024; W. Tang et al., 2018), but large group effects were reported in two studies both investigating tone language speaking amusics (Liu et al., 2012; Shao et al., 2016). For emotion recognition in full sentences, negligible to small effect sizes were reported in Lolli et al. (2015) and Pralus et al. (2019), but a large between-group difference was reported in Thompson et al. (2012). Note that with the number of participants from the present study (n = 18 amusics, n = 19 controls), we reached a power of 0.40–0.50 (computations performed with G*Power 3.1) for detecting a group main effect of medium size (f = 0.25) in a repeated-measure ANOVA with a 2- to 7-level within-subject factor (assuming r = 0.5 for the correlation among the repeated measures). We reached a power of 0.42 to detect a group difference of medium size (d = 0.5), with a Mann-Whitney one-sided test. We reached a power of 0.78–0.88 for detecting a group main effect of large size (f = 0.4) in a repeated-measure ANOVA with a 2- to 7- level within-subject factor (assuming r = 0.5 for the correlation among the repeated measures). We reached a power of 0.75 to detect a group difference of large size (d = 0.8) with a Mann-Whitney one-sided test. Power analyses thus highlight that we could reasonably expect to detect large group effect sizes in our data, and medium effect sizes to some extent but not small ones. This would mostly affect the EMO, AS, Audimots, Switchipido, and SIGL tasks. Note however that comparisons between studies remain tentative as most of the tasks differed in the details of their implementation from the tasks and materials used here.

3. Specific material and results

3.1. Auditory tests

3.1.1. Pitch tasks (PCD, DCI, STM)

3.1.1.1. Material. Pitch change detection (PCD)

On each trial, participants were presented with a sequence of five tones. On some trials (standard), all five tones had identical pitch, while on other trials (deviant), the pitch of the fourth tone was different (see Fig. 1A), as in Pralus et al. (2021). The standard fundamental frequency varied between trials and could take equiprobably the following values: 165, 196, 262, or 392 Hz. For the different trials, the changes relative to the standard tone were 0.125, 0.25, 0.5, 1, 2, or 4 semitones, either up or down compared to the standard. The deviant fundamental frequencies were between 131 and 494 Hz. Participants had to indicate whether the fourth tone was changed in comparison to the other tones by selecting "identical" or "different" on the screen after the end of the five tone sequences, as in Albouy et al. (2015a) and Hyde and Peretz (2004). 64 sequences were presented, with 16 identical trials and 48 different trials (one trial per deviant size, per direction of change: up or down, and per standard fundamental frequency).

Pitch direction of change identification (DCI)

On each trial, participants listened to two tones at different frequencies (see Fig. 1B). The fundamental frequencies of the tones were between 123 and 523 Hz. The difference between the two tones could be equiprobably 1, 2, 3, 4, 5, 6, or 7 semitones, either up or down, as in Pralus et al. (2021). Participants had to determine if the second tone was higher (up) or lower (down) in pitch by selecting "rising" or "falling" on the screen ("monte" or "descend" in French) after the end of the second

tone. There were 56 sequences of two tones, with 28 up and 28 down sequences.

Pitch short-term memory of four tone-sequences (STM)

On each trial, participants were presented with two melodies of four tones (S1 and S2) separated by a silent delay of 1000ms, as in Pralus et al. (2021). Each melody lasted 2300ms. The two sequences could be either identical or different (see Fig. 1C). The fundamental frequencies of the harmonic tones were between 262 and 440 Hz (corresponding to notes C4, D4, E4, F4, G4, A4, all belonging to the C-major scale). Identical tones were not repeated consecutively in a sequence, and all sequences entailed a change of contour. For the different trials, changes in S2 could occur on the second or third tone with changes of 3, 4, 5, 7, or 9 semitones relative to the tone at the same position in S1, always creating a change of melodic contour between S1 and S2. 32 trials were presented, with 16 identical trials and 16 different trials. Participants had to indicate whether S2 was identical or different from S1 by selecting "identical" or "different" on the screen after the end of S2.

3.1.1.2. Results. The performance of the two participant groups in the three pitch tasks (PCD, DCI, STM) is illustrated in Fig. 1D. We present first an analysis of the overall percentages of correct responses for the three tasks, and then a detailed analysis for each of the three tasks, assessing the different types of trials (identical vs. different trials for PCD and STM, and the effect of the size of pitch changes for all three tasks). Graphs illustrating the effects of type of trials and/or difficulty for each task are provided in the supplemental material.

PCD, DCI, and STM results

Percentages of correct responses for the three pitch-specific tasks (averaged across all trial types for each task) were analyzed with a repeated-measure ANOVA. A 2 × 3 ANOVA was performed with Group (amusics, controls) as a between-participants factor and Task (PCD, DCI, STM) as a within-participant factor. Statistical analyses revealed a main effect of Group [$F(1,35) = 29.91$, $p < 0.001$, $\eta_p^2 = 0.43$, Cohen's $f = 0.88$] and Task [$F(1.56,54.72) = 28.70$, $p < 0.001$, $\eta_p^2 = 0.45$, Cohen's $f = 0.87$]. Participants with amusia exhibited lower performance than controls. Post-hoc tests showed that participants had higher performance for PCD and STM than for DCI (all $p < 0.001$). Additionally, performance was significantly lower for PCD than for STM ($p = 0.04$). A marginal interaction between Group and Task was observed [$F(1.56,54.72) = 3.26$, $p = 0.06$, $\eta_p^2 = 0.08$, Cohen's $f = 0.25$] (Fig. 1D).

PCD results

Percentages of correct responses for "different" trials were analyzed with a 2 × 6 ANOVA with Group (amusics, controls) as a between-participants factor and Difficulty (pitch interval between the different tone and the standard tones: 0.125, 0.25, 0.5, 1, 2, 4 semitones) as a within-participant factor. The main effect of Group [$F(1,35) = 36.87$, $p < 0.001$, $\eta_p^2 = 0.51$, Cohen's $f = 0.98$] and the main effect of Difficulty [$F(2.9,101.46) = 65.26$, $p < 0.001$, $\eta_p^2 = 0.65$, Cohen's $f = 1.33$] were significant, as was their interaction Group*Difficulty [$F(2.9,101.46) = 20.93$, $p < 0.001$, $\eta_p^2 = 0.37$, Cohen's $f = 0.74$]. Post-hoc tests showed that performance of amusic participants decreased for pitch interval sizes of 0.5, 0.25, and 0.125 semitone relative to pitch interval sizes of 1, 2, and 4 semitones (all $p < 0.05$, excepted between 0.5 and 1: $p = 1$). For control participants, their performance decreased only for 0.125 semitone pitch interval compared to all other pitch interval sizes (all $p < 0.001$). Significant differences were observed between participants with amusia and controls for the 0.125, 0.25, and 0.5 semitone pitch interval sizes (all $p < 0.05$). Post-hoc tests revealed no significant differences between amusic and control participants for the 1, 2, and 4 semitones pitch interval sizes (all $p > 0.77$) (Supplementary Fig. S1A).

For the "identical" trials, we analyzed the percentage of correct responses and compared them between groups with a Mann-Whitney test because data were not normally distributed in both groups (Shapiro-Wilk tests, all $p < 0.05$). A significant difference was observed between participants with amusia and controls, with more false alarms in amusic participants ($W = 133$; $p = 0.04$, r_biserial = 0.33, Cohen's $d = 0.69$) (Supplementary Fig. S1A).

In addition, response times for correct responses on "different" trials were analyzed with 2 × 4 ANOVA with Group (amusics, controls) as a between-participants factor and Difficulty (0.5, 1, 2, 4 semitones) as a within-participant factor. As eight amusic participants and one control had 0% of correct responses for the 0.125 and 0.25 semitone pitch changes, these difficulty levels were removed from the response times analysis. The main effect of Group was significant but with a small effect size [$F(1,35 = 3.78$, $p = 0.06$, $\eta_p^2 = 0.1$, Cohen's $f = 0.27$]. The main effect of difficulty was not significant [$F(2.41,84.49) = 2.20$, $p = 0.11$, $\eta_p^2 = 0.06$, Cohen's $f = 0.18$]. The interaction Group*Difficulty was not significant either [$F(2.41,84.49) = 0.43$, $p = 0.69$, $\eta_p^2 = 0.01$, Cohen's $f = 0$]. Response times for the identical trials were analyzed and revealed a significant effect of the group ($t(35) = 2.06$, $p = 0.047$, Cohen's $d = 0.68$). Amusic participants had longer response times than the control participants both for different and identical trials (Supplementary Fig. S1B).

DCI results

Percentages of correct responses were analyzed with a 2 × 7 ANOVA with Group (amusics, controls) as a between-participants factor and Difficulty (difference between the two pitches: 1, 2, 3, 4, 5, 6, or 7 semitones) as a within-participant factor. The main effect of Group [$F(1,35) = 11.20$, $p = 0.002$, $\eta_p^2 = 0.24$, Cohen's $f = 0.52$] was significant, we observed that participants with amusia had lower performance than did controls. The main effect of Difficulty [$F(3.87,135.51) = 16.33$, $p < 0.001$, $\eta_p^2 = 0.32$, Cohen's $f = 0.65$] was significant. Participants had decreased performance when the pitch interval size was 1, 2, 3

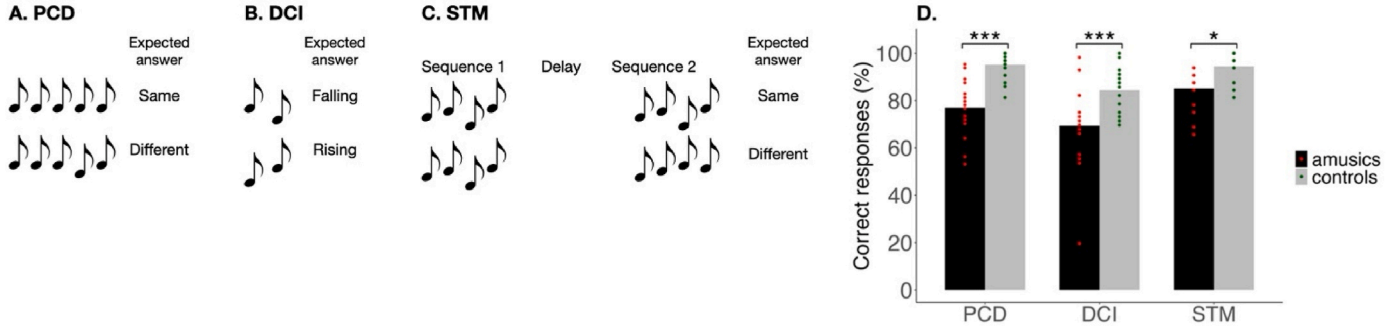


Fig. 1. Pitch tasks. **A**. In the pitch change detection (PCD) task, participants have to determine whether the fourth tone is identical or different from the others. **B**. In the pitch direction of change identification (DCI) task, they have to determine if the pitch contour is "falling" or "rising" from the first to the second tone. **C**. In the Short-Term Memory (STM) task, participants have to determine if the second sequence is identical or different from the first one. **D**. Mean percentages of correct responses of amusic (black bars) and control (gray bars) participants for the PCD, DCI, and STM tasks. Dots represent individual data for amusic (red) and control (green) participants. Between-group differences: *p < 0.05; **p < 0.01; ***p < 0.001.

semitones compared to 4, 5, 6, and 7 semitones (all $p < 0.05$, excepted between, 2 and 5 semitones: $p = 0.35$, 3 and 5 semitones: $p = 0.55$). In addition, we observed lower performances for the 1 pitch interval size compared to the 2 and 3 pitch interval sizes (all $p < 0.05$). Then, participants had lower performances for the 5-semitones pitch interval size compared to the 7 one ($p = 0.007$). No significant interaction between Group and Difficulty was observed [$F(3.87,135.51) = 1.05$, $p = 0.38$, $\eta_p^2 = 0.03$, Cohen's $f = 0.04$] (Supplementary Fig. S2A).

In addition, response times for correct responses were analyzed with a 2 × 7 ANOVA with Group (amusics, controls) as a between-participants factor and Difficulty (1, 2, 3, 4, 5, 6, 7 semitones) as a within-participant factor. The data from one amusic participant who had 0% correct responses for two levels of difficulty were excluded for this response time analysis. The main effect of Group was not significant [$F(1,34) = 0.58$, $p = 0.45$, $\eta_p^2 = 0.02$, Cohen's $f = 0$]. The main effect of difficulty was significant [$F(4.51,153.48) = 2.41$, $p = 0.04$, $\eta_p^2 = 0.07$, Cohen's $f = 0.2$]. Post-hoc tests revealed that participants had faster RT when the pitch interval size was 6 and 7 semitones compared to 1 semitone (all $p < 0.05$). Participants' response times were not significantly different between all the other conditions (all $p > 0.8$). The interaction between Group and Difficulty was not significant [$F(4.51,153.48) = 0.83$, $p = 0.52$, $\eta_p^2 = 0.02$, Cohen's $f = 0$] (Supplementary Fig. S2B).

STM results

Percentages of correct responses for "different" trials were analyzed with a 2 × 2 ANOVA with Group (amusics, controls) as a between-participants factor and Difficulty (size of the pitch interval between the original tone in S1 and the changed tone in S2: small, large changes) as a within-participant factor. 3 and 4-semitones pitch interval sizes were considered as small changes (8 trials), and 5, 7, and 9-semitones as large changes (8 trials) 3, 4, 5, 7, or 9 semitones. The main effect of Group [$F(1,35) = 10.2$, $p = 0.003$, $\eta_p^2 = 0.23$, Cohen's $f = 0.5$] was significant, with lower performance in the amusic group compared to the control group. The main effect of Difficulty [$F(1,35) = 39.55$, $p < 0.001$, $\eta_p^2 = 0.53$, Cohen's $f = 1.02$] was significant. We observed that participants had higher performance when the pitch change was large (Supplementary Fig. S3A). No significant interaction was observed between Group and Difficulty [$F(1,35) = 1.88$, $p = 0.18$, $\eta_p^2 = 0.05$, Cohen's $f = 0.15$].

For the "identical" trials, we compared the percentages of correct responses between groups with a Mann-Whitney test because data were not normally distributed in both groups (Shapiro-Wilk tests, all $p < 0.05$). No significant difference was observed between amusic and control participants ($W = 150$; $p = 0.46$, r_biserial $= -0.12$, Cohen's $d = 0.24$) (Supplementary Fig. S3A).

In addition, we analyzed response times for correct responses for "different" trials with 2 × 2 ANOVA with Group (amusics, controls) as a between-participants factor and Difficulty (small, large changes) as a within-participant factor. A marginal effect of Group was found [$F(1,35 = 3.99$, $p = 0.054$, $\eta_p^2 = 0.10$, Cohen's $f = 0.28$]. Amusic participants tended to have longer RTs than control participants (Supplementary Fig. S3B). The main effect of Difficulty was significant [$F(1,35) = 5.33$, $p = 0.03$, $\eta_p^2 = 0.13$, Cohen's $f = 0.34$], revealing that participants had shorter RTs for the large pitch changes than for the small pitch changes. The interaction Group*Difficulty was not significant [$F(1,35) = 0.09$, $p = 0.77$, $\eta_p^2 = 0.002$, Cohen's $f = 0$].

Response times for correct responses for the "identical" trials were compared between groups and revealed a significant effect of Group ($t(35) = 2.87$, $p = 0.007$, Cohen's $d = 0.95$). For "identical" trials participants with amusia had longer response times than the control group (Supplementary Fig. S3B).

3.1.2. Pitch short-term memory with six tone-sequences (STM_6)

3.1.2.1. Material. Only thirteen participants with amusia and six control participants out of the 19 participants of each group performed this task during the main testing session after the STM task. We completed the dataset for this task using data from four individuals with amusia and twelve matched controls from another study (Hoarau et al., in prep). Thus, in total we had data from 17 amusic and 18 control participants for the STM_6 task (see Supplementary Table S1 for demographic data). We also retrieved data from the STM task with four-tone sequences for the same additional participants to compare directly performance in STM and STM_6 for these 35 participants.

The test procedure was the same as for STM with four tones except that here participants were presented with two melodies of six tones. Each melody lasted 3500ms with a delay of 1000ms between S1 and S2. The fundamental frequencies of the harmonic tones were between 262 and 523 Hz (corresponding to notes C4, D4, E4, F4, G4, A4, B4, C5, C-major scale). For the different trials, changes could occur on the second, third, fourth, or fifth tones with changes of 3, 5, 5, 7, 9, or 10 semitones, always creating a change of melodic contour between S1 and S2. 32 trials were presented, including 16 identical and 16 different trials.

3.1.2.2. Results. Analyses of the STM_6 task and comparisons between STM and STM_6 were carried out separately from the initial PCD, DCI, and STM analyses and are presented as supplementary material (after the demographic information in Table S1) as they were not performed on the same participants. For STM_6, for the "different" trials, we found a main effect of Group and Difficulty for the percentages of correct responses, with amusic participants showing lower performance than controls, as for STM, however only an effect of Difficulty was observed on response times. For the 'same' trials, as for STM, no difference between amusic and control participants was observed for the percentages of correct responses, but response times were longer for amusic participants than for control participants. The comparison between STM and STM_6 revealed the expected effect of the length of the sequence (lower percentages of correct responses and slower response times with six-tone sequences than four-tone sequences).

3.1.3. Auditory Stream segregation (AS)

3.1.3.1. Material. Participants were presented with a sequence of tones composed of ABA triplets (see Fig. 2A) as in Pralus et al. (2021). The tone A was the standard tone and B a tone with varying frequency. The duration of all tones was 100ms. The ISI between A and B was 20ms, and the interval between two ABA triplets was 140ms. The frequency of A was 196 Hz, and the frequency of B took the following values, in ascending or descending order: 196, 247, 294, 440, 659, or 988 Hz. Five ABA triplets were repeated for each frequency of B. The sequence started with a fundamental frequency of B at 440 Hz, going down to 196 Hz, then up to 988 Hz, and down again to 196 Hz. This pattern was repeated five times (Fig. 2B). The sequence ended with a frequency of B at 294 Hz. During the sequence, participants had to tell if they perceived one stream (meaning they perceived the sequence as "integrated") or two streams (meaning they perceived the sequence as "segregated"). They gave their answer by selecting either "1 stream" or "2 streams" on the screen. Once one button was selected, it remained selected until the participant changed their answer. Participants could respond as many times as they wanted during the sequence. The duration of the sequence was 2.5 min.

3.1.3.2. Results. We analyzed the time spent in one-stream and two-stream percepts, the mean frequency of B at the change of percept from one to two streams, and the number of percept changes.

The total time spent in one-stream or two-stream percepts was analyzed with a 2 × 2 ANOVA with Group (amusics, controls) as a between-participants factor and Percept (one or two streams) as a within-participant factor. Analyses revealed no significant effect of Group [$F(1,35) = 1.51$, $p = 0.23$, $\eta_p^2 = 0.04$, Cohen's $f = 0.12$]. The main

effect of Percept was significant [F (1,35) = 84.51, p < 0.001, $\eta_p^2 = 0.71$, Cohen's f = 1.5] revealing that participants spent more time in the two-stream percept than in the one-stream percept. The interaction Group*Percept was not significant [F (1,35) = 0.02, p = 0.89, $\eta_p^2 = 0.0004$, Cohen's f = 0] (Fig. 2C).

The mean frequency at the change of percept from one to two streams and the number of percept changes were compared between groups with a Mann-Whitney test because data were not normally distributed within each group (Shapiro-Wilk tests, all p < 0.05). The mean frequency at the change of percept from one to two streams was not significantly different between amusic and control participants (W = 165, p = 0.86, r_biserial = −0.03, Cohen's d = −0.06) (Fig. 2D). No significant difference was observed for the number of percept changes (mean score controls = 7.58 (±1.64); mean score amusics = 8.21 (±2.88); W = 177.5, p = 0.83, r_biserial = 0.04, Cohen's d = 0.08).

3.1.4. Emotion recognition in spoken sentences (EMO)

3.1.4.1. Material. In the emotion recognition task, participants listened to 20 semantically neutral French sentences from Pralus et al. (2019, 2021). These sentences were: "L'avion est presque plein." ("The plane is almost full.") and "J'espère qu'il va m'appeler bientôt." ("I hope he will call me soon."). They were pronounced with different emotions by female or male talkers. For each emotion (joy, sadness, anger, fear, neutral), two tokens were used for each sentence, half pronounced by a male voice and half by a female voice. The average duration of the stimuli was 1470ms (±278ms). For each trial, after the end of the sentence, participants had to select on the screen the recognized emotion first, and then rate the intensity of the selected emotion from 1 (not intense) to 5 (very intense) except for stimuli judged as neutral (Fig. 3A).

3.1.4.2. Results. For emotion recognition in full sentences, we analyzed the percentage of correct responses with a 2 × 5 ANOVA with Group (amusics, controls) as a between-participants factor and Emotion (Joy, Sadness, Anger, Fear, Neutral) as a within-participant factor. The ANOVA revealed no significant effect of Group [F (1,35) = 0.32, p = 0.57, $\eta_p^2 = 0.01$, Cohen's f = 0]. The main effect of Emotion was significant [F (2.58,90.37) = 5.94, p = 0.002, $\eta_p^2 = 0.15$, Cohen's f = 0.37]. Post-hoc tests revealed decreased performance for Fear and Neutral compared to other emotions (all p < 0.05). No significant interaction was observed between Group and Emotion [F (2.58,90.37) = 0.49, p = 0.66, $\eta_p^2 = 0.01$, Cohen's f = 0] (Fig. 3B).

For intensity ratings (of correct recognitions), we performed a 2 × 4 ANOVA with Group (amusics, controls) as a between-participants factor and Emotion (Joy, Sadness, Anger, Fear) as a within-participant factor. A marginal effect of Group was found, with a small effect size [F (1,35) = 3.98, p = 0.054, $\eta_p^2 = 0.10$, Cohen's f = 0.28]. The main effect of Emotion was significant [F (3,105) = 15.54, p < 0.001, $\eta_p^2 = 0.31$, Cohen's f = 0.63]. Post-hoc tests showed that Joy and Anger were rated as being more intense than Fear and Sadness (all p < 0.01). No significant interaction was observed [F (3,105) = 1.36, p = 0.26, $\eta_p^2 = 0.04$, Cohen's f = 0.1] (Fig. 3C).

3.1.5. Emotion recognition in spoken vowels (EMO_v)

3.1.5.1. Material. Only fourteen participants with amusia and six control participants performed the task in the same testing session (same participants as for the STM_6 task, the dataset was completed with the same additional participants as for STM_6, see Supplementary Table S1 for demographic data). We also retrieved data from the EMO task with full sentences for the same participants to compare directly EMO and EMO_v. The test procedure was the same as the EMO test above except that here the stimuli used were/a/vowels all pronounced by women's voices (stimuli from Charpentier et al., 2018, see Pralus et al., 2019 for the full experimental procedure). The duration of the vowel stimuli was 400ms, and there were twenty trials (four for each of the following emotions: joy, sadness, anger, fear, neutral).

3.1.5.2. Results. Analyses of the EMO_v task and comparisons between EMO and EMO_v were carried out separately from the initial EMO task

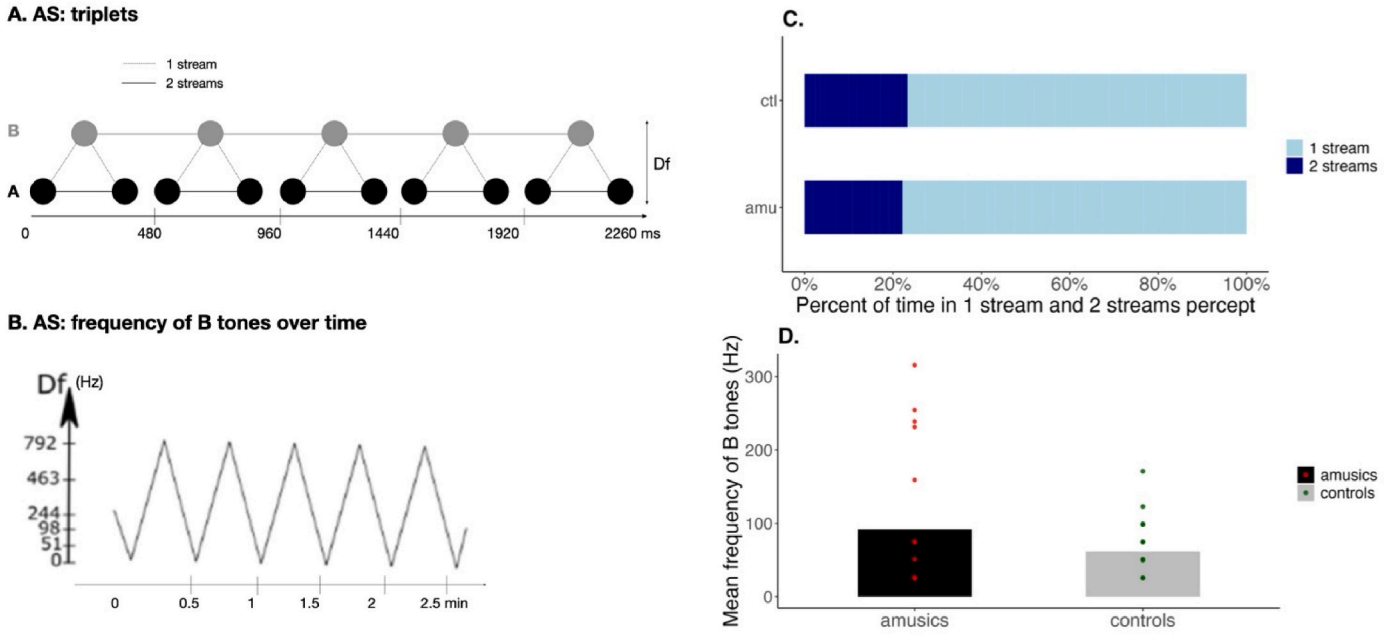


Fig. 2. A. In the Auditory Stream segregation (AS) task, participants hear a sequence of ABA triplets. The tone A has a fixed frequency whereas the frequency of B changes across time (Df: frequency difference between A and B tones). Participants have to determine if the sequence is perceived as one stream or two streams. **B.** Change of tone B frequency relative to tone A frequency across time (Df), see main text for details. **C.** Percent of time spent in the one stream (dark blue) and two streams (light blue) percepts, on average, for each participant group (amusics and controls). **D.** Mean frequency of B tones (Hz) at the change of percept from one to two streams of amusic (black bars) and control (gray bars) participants. Dots represent individual data for amusic (red) and control (green) participants. No significant between-group differences were observed.

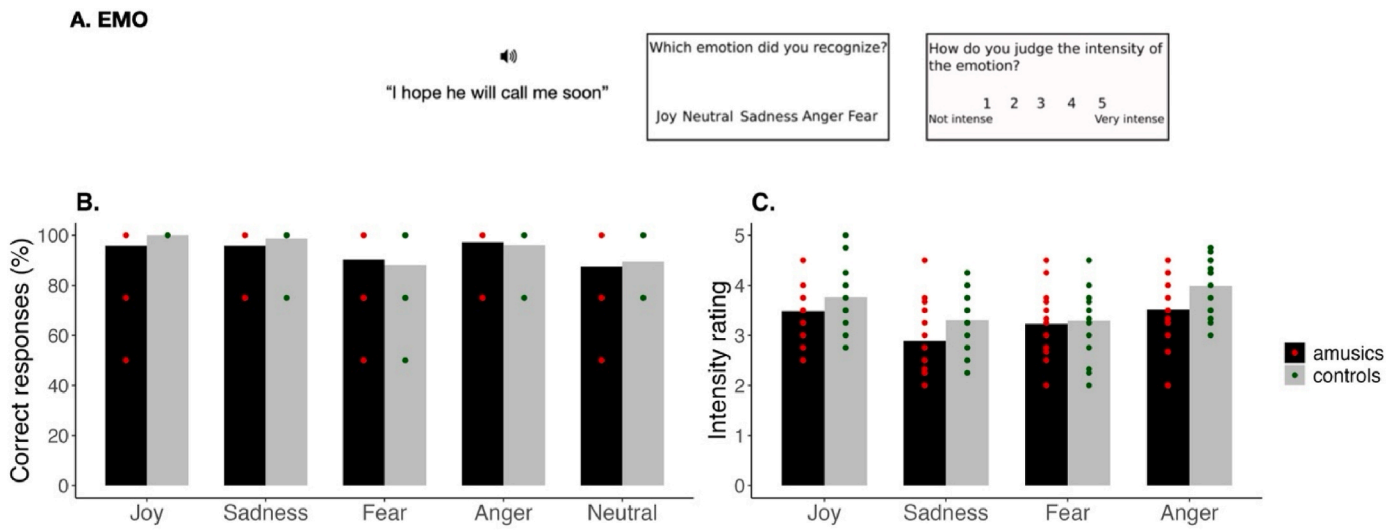


Fig. 3. A. In the emotion recognition tasks (EMO) participants hear a sentence pronounced with different emotions (joy, sadness, anger, fear, neutral) and have to select the recognized emotion and rate the intensity of the selected emotion (except for neutral emotion). **B.** Percentages of correct responses for emotion recognition for each emotion on average in each group: amusics (dark bars) and controls (gray bars). **C.** Mean intensity rating for each emotion for amusics (dark bars) and controls (gray bars). Dots represent individual data (amusics in red and controls in green). Between-group differences: *p < 0.05; **p < 0.01; ***p < 0.001.

and are presented as supplementary material since they were not performed on the same participants. For EMO_v, we found a main effect of the group for recognition scores, with participants with amusia showing lower performance than controls.

3.1.6. *Speech-in-noise perception (Audimots)*

3.1.6.1. Material. For each trial in Audimots (Moulin et al., 2013), participants heard a word in noise and had to select the correct answer among four propositions on the screen (Fig. 4A). The noise was presented continuously during the block, and the word lists appeared on the screen 800ms prior to the presentation of the target word. The three foils could be dissimilar to the target word phonetically (easy trials) or similar, with differences only on vowels (difficult trials) or differences only on the initial consonants (very difficult trials). Two types of noise of the same long-term spectra were used, a speech noise and cocktail party noise. The Cocktail Party noise corresponds to a multi-speaker noise (16 speakers: 8 females, 8 males). The speech noise is a background noise that shares the spectral and acoustic information of the speech signal but where the speech information has been removed, i.e., random noise of the same frequency spectrum as the cocktail party noise (Bourgeois–Vionnet et al., 2020). For all participants, a signal-to-noise ratio (SNR) of −6 dB was used with noise level fixed at 65 dB SPL A, measured at the participant's ear using a Brüel & Kjaer type 2239 sonometer. There were 60 trials for each noise type, 20 easy trials, 20 difficult trials, and 20 very difficult trials.

3.1.6.2. Results. Percentages of correct responses were computed for the two types of noise and the three phonological difficulty levels and analyzed with a 2 × 2 × 3 ANOVA with Group (amusics, controls) as a between-participants factor, Noise (SpeechNoise, CocktailParty) and Difficulty (Easy, Difficult, Very Difficult) as within-participant factors was performed. The main effect of Group was not significant [$F(1,33) = 0.22$, $p = 0.64$, $\eta_p^2 = 0.007$, Cohen's $f = 0$]. The main effect of Noise was significant [$F(1,33) = 25.79$, $p < 0.001$, $\eta_p^2 = 0.44$, Cohen's $f = 0.84$]. Participants reached higher performance in the Speech Noise condition than in Cocktail Party. The main effect of Difficulty was significant [$F(2,66) = 103.53$, $p < 0.001$, $\eta_p^2 = 0.76$, Cohen's $f = 1.72$]. Participants reached higher performance for the easy trials, when the three foils were phonetically very dissimilar from the target word, than for the difficult and very difficult trials with phonological neighbors (all $p < 0.001$). Moreover, participants had lower performance for the very difficult trials, when the differences between the four propositions were on the initial consonant, than for the difficult trials where the differences were on the vowels ($p < 0.001$). No significant interaction was observed (all $p > 0.07$) (Fig. 4B). Data from one amusic participant and one control participant were excluded from Audimots analyses due to a technical error, the SNR used during testing (+4 dB) was different from the SNR used for the other participants (-6 dB).

Response times were analyzed with a 2 × 2 × 3 ANOVA with Group (amusics, controls) as a between-participants factor, Noise (SpeechNoise, CocktailParty) and Difficulty (Easy, Difficult, Very Difficult) as within-participants factors. The main effects of Group [$F(1,33) = 0.92$, $p = 0.34$, $\eta_p^2 = 0.03$, Cohen's $f = 0$] and Noise [$F(1,33) = 0.04$, $p = 0.83$, $\eta_p^2 = 0.001$, Cohen's $f = 0$] were not significant. The main effect of Difficulty was significant [$F(1.64,53.99) = 12.77$, $p < 0.001$, $\eta_p^2 = 0.28$, Cohen's $f = 0.58$]. Participants had shorter response times for the easy and difficult trials than for the very difficult trials (all $p < 0.001$). No significant response time difference was observed between easy and difficult trials ($p = 0.17$). The ANOVA revealed no significant interactions (all $p > 0.53$) (Supplementary Fig. S6).

3.1.7. *Auditory questionnaires*

3.1.7.1. Material. Participants were asked to fill out two paper questionnaires about hearing quality and listening effort: the 15iSSQ and the EEAS. The 15iSSQ is a short form of Gatehouse and Noble (2004)'s Speech, Spatial, and Qualities of hearing scale, which is one of the most widely used self-report measures for hearing abilities (see Moulin et al., 2015, 2019 for the French translation used here). Participants had to rate the answer to each question on a scale that ranges from 0 to 10 (see Fig. 5A): 10 means they are capable of doing what is described in the corresponding question whereas 0 means that they cannot do what is described. The items are grouped into three sub-scales: speech perception (5 questions focusing on speech-in-noise), spatial hearing (5 questions), and qualities of hearing (5 questions). The mean scores for each of the three sub-scales were calculated. Then participants completed a listening effort questionnaire, the EEAS (Ferschneider and Moulin, 2023), which was adapted from the Effort Assessment Scale (EAS, see Alhanbali et al., 2017). The adapted version of EAS (EEAS = Extended EAS) is composed of ten items. Participants had to fill out the questionnaire by putting a mark on a scale that ranges from 0 (no effort) to 10

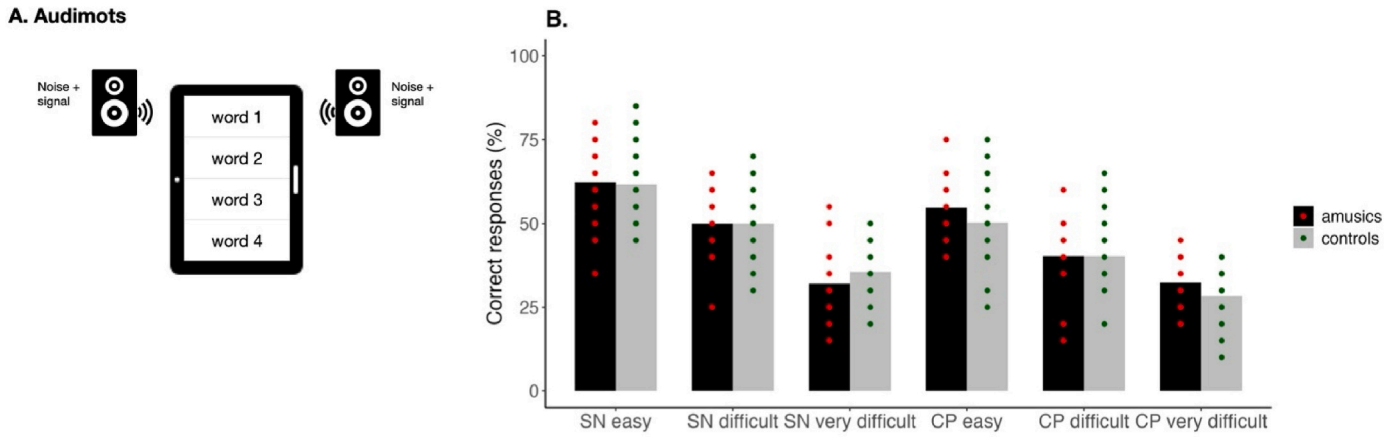


Fig. 4. **A.** In the speech-in-noise Audimots test, participants hear a word in noise and have to select the correct answer among four propositions on the screen. **B.** Mean percentage of correct responses for amusics (dark bars) and controls (gray bars). Dots represent individual data (amusics in red and controls in green). There are two types of noise: SN = SpeechNoise and CP = CocktailParty. The four words propositions could be dissimilar to the target word phonetically (easy trials) or similar, with differences only on vowels (difficult trials) or differences only on the initial consonants (very difficult trials). No significant between-group differences were observed.

(lots of effort, see Fig. 5B). The items are grouped in three sub-scales: one sub-scale on hearing in silence (3 questions), two sub-scales on hearing in noise: one with questions matched to the hearing in silence questions (3 questions) and one containing question specifically related to hearing in noise (4 questions). The mean scores for each of the three sub-scales were calculated.

3.1.7.2. Results. Scores of the two hearing questionnaires were analyzed for each sub-scale (as in Moulin and Richard, 2016).

For the 15iSSQ questionnaire, one-tailed Student t tests with Bonferroni correction were performed on mean ratings across questions for each subscale (Speech Perception, Spatial Hearing, Qualities of Hearing). No difference between amusic and control participants were observed for the Speech Perception (t (35) = -0.87, p = 0.59, Cohen's d = −0.29) and Spatial Hearing sub-scales (t (35) = -0.19, p = 1, Cohen's d = −0.06). For the Qualities of Hearing sub-scale, a significant group effect was observed (t (35) = -3.54, p = 0.002, Cohen's d = −1.17). Participants with amusia reported lower qualities of hearing scores compared to controls (Fig. 5C). One-tailed Mann Whitney tests (data were not normally distributed, $p < 0.05$) corrected with Bonferroni

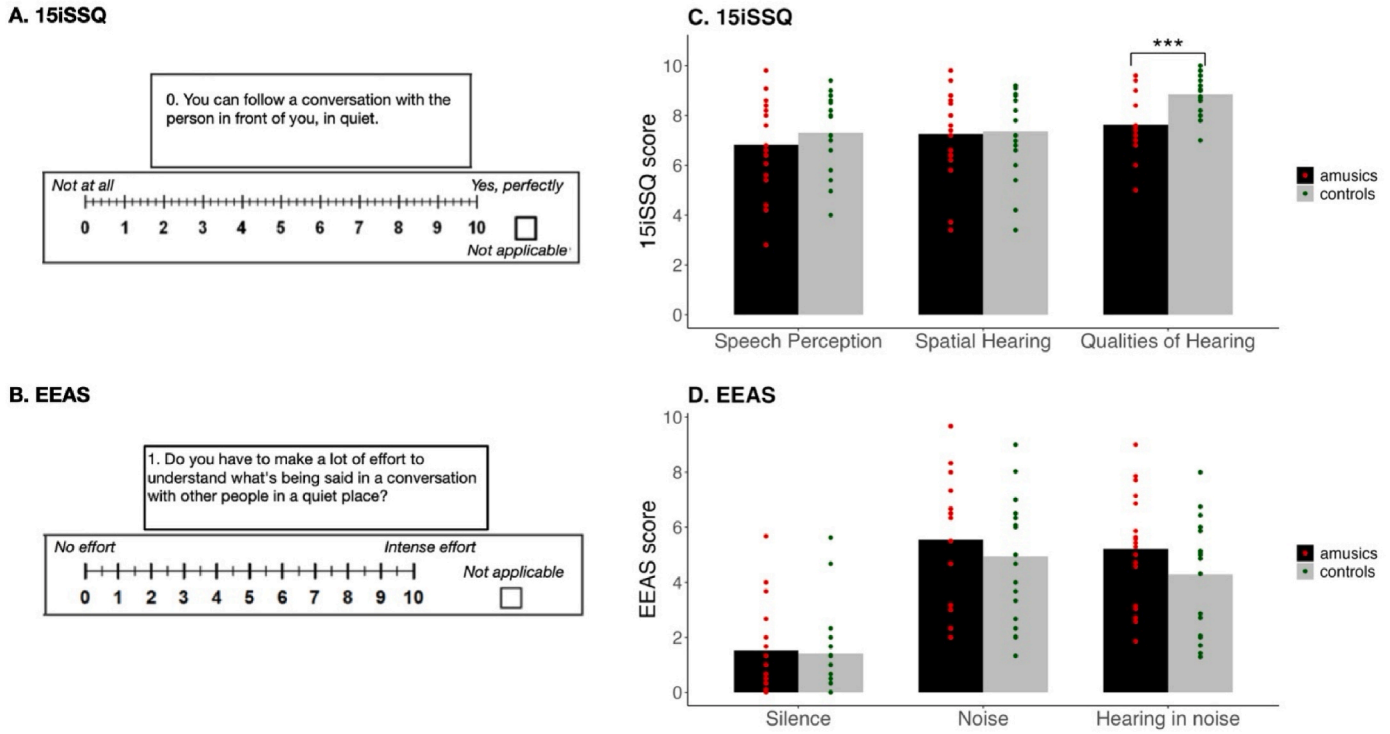


Fig. 5. 15iSSQ and EEAS questionnaires. **A.** For the 15iSSQ questionnaire, participants have to rate the answer to each question on a scale that ranges from 0 (they cannot do what is described in the corresponding question) to 10 (they are capable of doing what is described). **B.** For the EEAS questionnaire, they have to fill the questionnaire by putting a mark on a scale that ranges from 0 (no effort) to 10 (intense effort). **C.** Mean ratings for the three 15iSSQ sub-scales for amusic (black bars) and control (gray bars) participants. **D.** Mean ratings for the three EEAS sub-scales for amusic (black bars) and control (gray bars) participants. In C and D, dots represent individual data for amusic (red) and control (green) participants. Between-group differences: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

procedure were computed on the five questions of the qualities of Hearing sub-scale. Significant differences were observed for the two questions on music: "Can you easily recognize the different pieces of music you know?" (mean score controls = 8.61 (±1.28); mean score amusics = 6.03 (±2.76); W = 77, p = 0.008, r_biserial = 0.55, Cohen's d = 1.32) and "When you listen to music, does it sound clear and natural to you?" (mean score controls = 9.39 (±0.89); mean score amusics = 7.06 (±2.01), W = 49.5, p < 0.001, r_biserial = −0.71, Cohen's d = −2). Participants with amusia had a lower score than controls for these two questions. No significant differences between amusic and control participants were observed for the questions "Can you easily recognize the different people you know, by the sound of their voices?", "Can you tell the difference between certain noises, such as a car versus a bus, or water boiling versus food frying in a pan?", and "Do the everyday noises that you hear easily sound clear and distinct (not garbled, not mixed)?" (all p > 0.22).

For the EEAS questionnaire, analyses were performed on mean ratings across questions for each subscale (Silence, Noise, Hearing in Noise). One-tailed Mann Whitney tests (data were not normally distributed, p < 0.05) corrected with Bonferroni procedure were used. No difference between amusic and control participants were observed for the Silence (W = 177, p = 1, r_biserial = 0.03, Cohen's d = 0.06), Noise (W = 195, p = 0.71, r_biserial = 0.14, Cohen's d = 0.29), and Hearing in Noise (W = 215.5, p = 0.27, r_biserial = 0.26, Cohen's d = 0.54), sub-scales (Fig. 5D).

Overall, our results revealed impaired performance for the three pitch tasks (PCD, DCI, STM) for amusic participants compared to control participants. For the EMO recognition task with sentences, as expected based on Pralus et al. (2019) using the same material, participants with amusia performed just as well as controls, but they were impaired for the EMO_v recognition task with vowels. For intensity ratings of EMO, the effect of group was marginally significant for sentences, yet with a small effect size, participants with amusia rated emotions as being less intense than controls did. However, amusic and control participants did not differ in intensity ratings in vowels. For the stream segregation task (AS), speech-in-noise task (Audimots), and listening effort questionnaire (EEAS), no differences were observed between amusic and control participants. In the 15iSSQ questionnaire, we observed a significant difference between amusic and control participants only for the two questions related to music.

3.2. Reading and copying tests

3.2.1. Reading test

3.2.1.1. Material. For the reading test (DeltaText, Vialatte et al., 2023), participants had to read a text aloud as rapidly and correctly as possible. The experimenter recorded reading time and number of errors. The text is grammatically correct but tells a somewhat nonsense story to avoid semantic anticipations.

3.2.1.2. Results. Group analyses were performed with a one-tailed Mann-Whitney test because data were not normally distributed in both groups (Shapiro-Wilk tests, all p < 0.05).

For the reading test (DeltaText), we analyzed the reading time (Fig. 6B) and the number of errors (Fig. 6C). The statistical analyses revealed no significant difference between amusic and control participants for reading time (W = 211.5, p = 0.11, r_biserial = 0.24, Cohen's d = 0.5) and number of errors (W = 194, p = 0.24, r_biserial = 0.13, Cohen's d = 0.26). In addition, the weighted speed index (number of correctly read words*max reading time/real reading time, Fig. 6D) was calculated and compared between groups (the maximum reading time was 90 s). No significant difference between groups was observed (W = 130, p = 0.11, r_biserial = −0.24, Cohen's d = −0.49).

3.2.2. Copying test

3.2.2.1. Material. For the copying test, "la Baleine Paresseuse" (meaning "the Lazy Whale", see (Duplat and Girier, 2006), participants had to copy a text hung on the wall as correctly as possible in 3 min. The text was placed in front of them at a distance of 50 cm and was written in a size font of 14. The experimenter counted the number of back-and-forth eye-movements between the text and the copying sheet and the number of characters correctly copied.

3.2.2.2. Results. Group analyses were performed with one-tailed Mann-Whitney tests because data were not normally distributed in both groups (Shapiro-Wilk tests, all p < 0.05).

We analyzed the number of characters copied correctly, the number of back-and-forth eye-movements between the text and the copying sheet, and the memory span (number of characters correctly copied/number of back-and-forth eye-movements). Data from one control participant was excluded from analyses of eye movements and span due to an error during testing, the number of back-and-forth eye-movements was not coded by the experimenter. The statistical test revealed no significant difference between amusic and control participants for the total number of characters correctly copied (W = 186, p = 0.68, r_biserial = 0.09, Cohen's d = 0.18 Fig. 7B). However, the number of back-and-forth eye-movements was significantly different between groups (W = 219.5, p = 0.03, r_biserial = 0.35, Cohen's d = 0.75, Fig. 7C). Participants with amusia did more back-and-forth eye-movements between the text and the copying sheet (mean: 34.17; SD: 10.92) than controls (mean: 27.39; SD: 8.89), and hence memorized fewer letters at a time. The analysis of this copying memory span revealed a significant effect of the group (W = 108.5, p = 0.047, r_biserial = −0.33, Cohen's d = −0.7), suggesting a smaller (poorer) memory span for the to-be-copied text information in amusic participants than in control participants (Fig. 7D).

3.3. Visuo-spatial tests

3.3.1. Switchipido arrow

3.3.1.1. Material. Global

For each trial, participants were presented with an image of a small arrow contained within a big arrow. One arrow was white and the other blue. Each arrow could be oriented downwards or upwards, and the two arrows could have the same or different orientations (Fig. 8A). Participants had to indicate the orientation of the big arrow, by selecting "Up" or "Down" on the screen, and ignore the orientation of the small one. 24 trials were presented.

Alternance

For each trial, participants were presented with a small arrow included in a big arrow, both included in a big circle. One arrow was white, the other was blue, and the circle was filled with the same color as the small arrow. Each arrow could be oriented downwards or upwards and the two arrows could have the same or different orientations (Fig. 8B). Participants had to indicate the orientation of the white arrow (which could thus randomly be the small or the big arrow), by selecting "Up" or "Down" on the screen, and ignore the orientation of the blue one. 32 trials were presented.

3.3.1.2. Results. We did not analyze the percentages of correct responses due to high performance (ceiling effect) in both groups. Reaction times of each visuo-spatial test were analyzed with Group (controls and amusics) and Condition as factors.

A 2 × 2 ANOVA was performed with Group (amusics, controls) as a between-participants factor and Condition (Global, Alternance) as a within-participant factor. No significant main effect of Group was observed [F (1,35) = 1.85, p = 0.18, η_p^2 = 0.05, Cohen's f = 0.15]. The

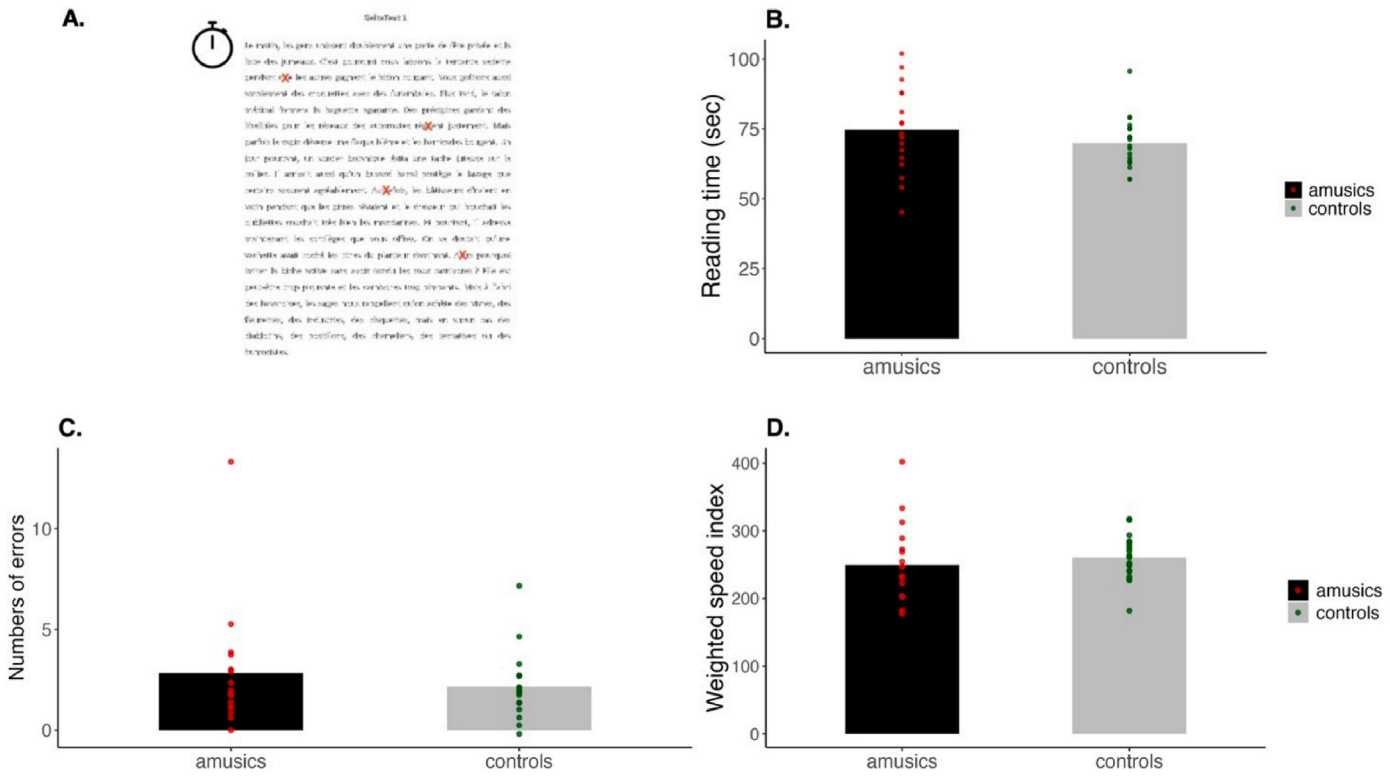


Fig. 6. Reading test (DeltaText). **A.** Example of DeltaText that participants had to read. Reading errors are crossed out (red cross). **B.** Reading time (in seconds). **C.** Number of reading errors. **D.** Results of the weighted speed index (number of correctly read words*max reading time/real reading time). In all three panels, average performance for the amusic group is indicated as a black bar and for the control group as a gray bar; dots represent individual data (amusics in red and controls in green). No significant between-group differences were observed.

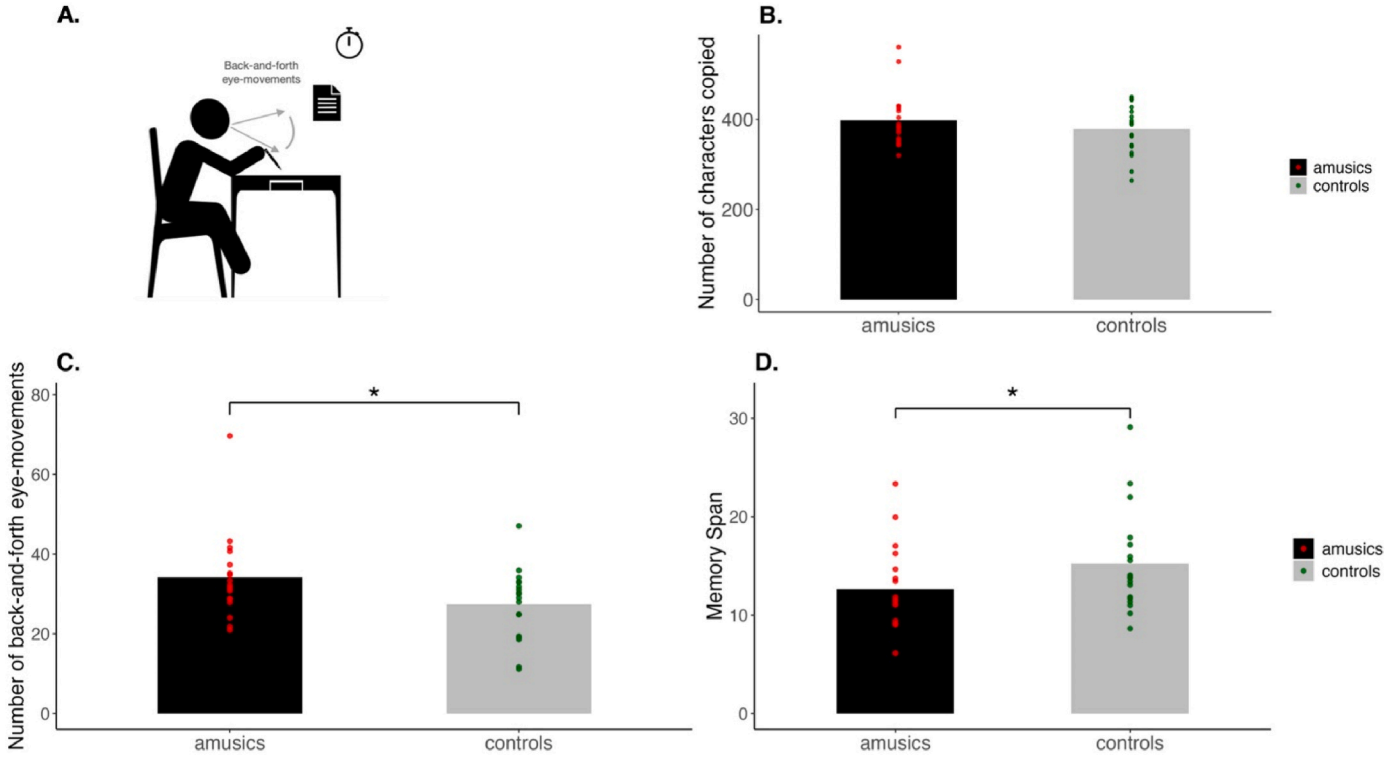


Fig. 7. Copying test ("la Baleine Paresseuse"). **A.** Schematic representation of the copying test experimental design. **B.** Number of characters correctly copied in 3 min. **C.** Number of back-and-forth eye-movements. **D.** Memory span (number of characters copied/number of back-and-forth eye-movements). In all three panels, average performance for the amusic group is indicated as a black bar and for the control group as a gray bar; dots represent individual data (amusics in red and controls in green). Between-group differences: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

main effect of Condition was significant [F (1,35) = 88.15, p < 0.001, η_p^2 = 0.72, Cohen's f = 1.53]. Participants had shorter reaction times for the Global condition than for the Alternance condition. No significant interaction was observed [F (1,35) = 1.48, p = 0.23, η_p^2 = 0.04, Cohen's f = 0.11] (Fig. 8F).

3.3.2. Switchipido triplet

3.3.2.1. Material. Global

Participants had to match a target stimulus presented at the top of the screen with one of two hierarchical symbols presented below, focusing on the global symbol and ignoring the small local symbols. The hierarchical symbols were global symbols made of small local symbols (different from the global one). Global and local symbols could represent a heart, a moon, a cross, a cup, or a star (Fig. 8C). After the appearance of the three stimuli on the screen, participants had to answer by clicking on one of the two hierarchical symbols. 27 trials were presented.

Simple

Participants had to match a target stimulus presented at the top of the screen with one of two hierarchical symbols presented below. Participants had to focus on the global symbols and the small local symbols composing the hierarchical symbols. The target stimulus could match with either of the two symbols, in either of the two levels of attention focalization (Fig. 8D). Participants were informed that there was no ambiguity, the target stimulus matched only with one of the two symbols. After the appearance of the three stimuli on the screen, participants had to answer by clicking on one of the two hierarchical symbols. 52 trials were presented.

Complex

In this task, the target presented at the top of the screen was also hierarchical, being made of smaller local symbols (different from the global one). Participants had to match the complex target presented at the top of the screen (by considering both the global target and the smaller local symbols composing it) with one of the two hierarchical symbols also focusing on the global symbols and the small local symbols (Fig. 8E). Either the global symbol or the smaller local symbol of the target could match with either local or global levels of the hierarchical symbols. After the appearance of the three stimuli on the screen, participants had to answer by clicking on one of the two large hierarchical symbols. 63 trials were presented.

3.3.2.2. Results.

We did not analyze the percentages of correct responses due to high performance (ceiling effect) in both groups. Reaction times of each visuo-spatial test were analyzed with Group (controls and amusics) and Condition as factors.

Data from one control participant was excluded from analyses due to a technical error, performance for the Complex condition was not saved. A 2 × 3 ANOVA was performed with Group (amusics, controls) as a between-participants factor and Condition (Global, Simple, Complex) as a within-participant factor. The main effect of Group was not significant [F (1,34) = 0.22, p = 0.64, η_p^2 = 0.006, Cohen's f = 0]. The main effect of Condition was significant [F (1.99,67.65) = 134.65, p < 0.001, η_p^2 = 0.79, Cohen's f = 1.94]. Post-hoc tests showed that the reaction time of participants was shorter for the Global condition than for the Simple and Complex conditions (all p < 0.001) and shorter for the Simple condition than for the Complex condition (p < 0.001). No significant interaction between Group and Condition was observed [F (1.99,67.65) = 0.47, p = 63, η_p^2 = 0.01, Cohen's f = 0] (Fig. 8G).

3.3.3. SIGL

3.3.3.1. Material. Local

The stimuli were large (global) hierarchical letters made of smaller (local) letters (different from the global one) presented at the top of the screen. Participants had to match the local letters (and ignore the global

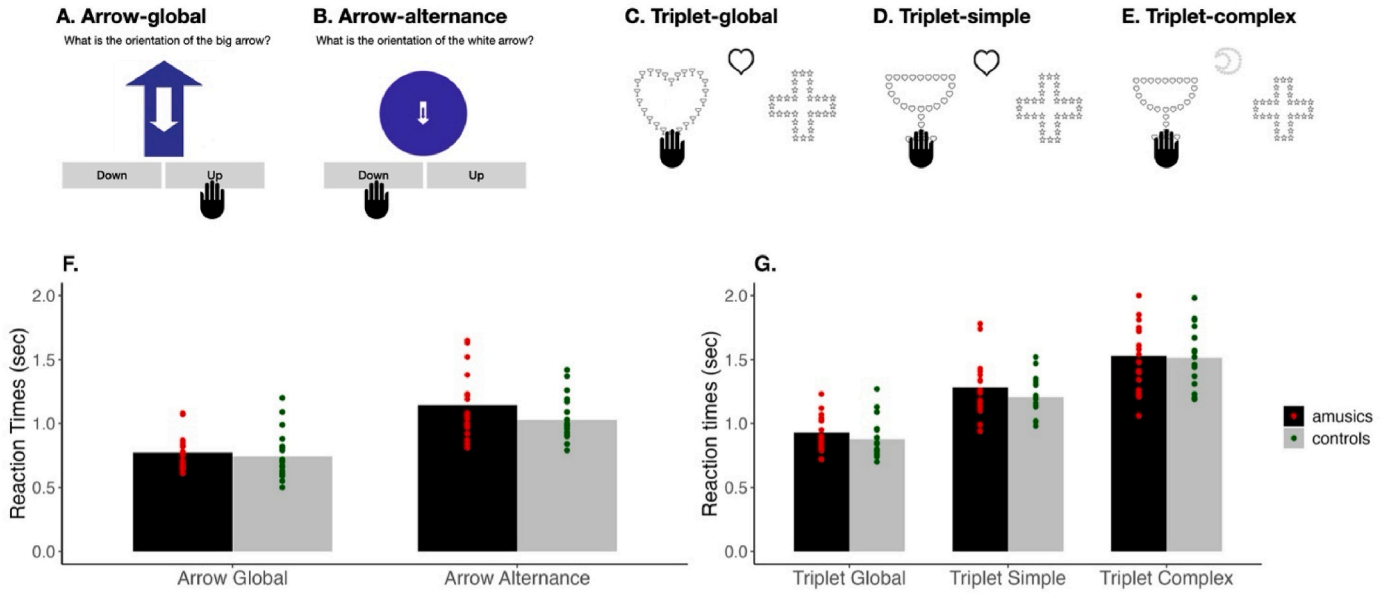


Fig. 8. Switchipido tasks. **A.** In the Arrow-global, participants have to indicate the orientation of the big arrow, by selecting "Up" or "Down" on the screen, and ignore the orientation of the small one. **B.** In Arrow-alternance, participants have to indicate the orientation of the white arrow (which could thus randomly be the small or the big arrow), by selecting "Up" or "Down" on the screen, and ignore the orientation of the blue one. **C.** In Triplet-global, they have to match a target stimulus presented at the top of the screen with one of two hierarchical symbols presented below, focusing on the global symbol and ignoring the small local symbols. **D.** In Triplet-simple, participants have to indicate which of the two symbols matched with the target stimulus by focusing on the global symbols and the small local symbols composing the hierarchical symbols. **E.** In Triplet-complex, participants have to match the complex target presented at the top of the screen (by focusing on the global target and the smaller local symbols composing it) with one of the two hierarchical symbols focusing on the global symbols or the small local symbols. **F.** Reaction times (in seconds) for Arrow tasks. **G.** Reaction times (in seconds) for Triplet tasks. In F and G, average performance for the amusic group is indicated as a black bar and for the control group as a gray bar; dots represent individual data (amusics in red and controls in green). No significant between-group differences were observed.

one) with one of the two letters presented at the bottom of the screen by selecting one of them. Local letters were E or M and global letters were H, T, or A (Fig. 9A). After the appearance of the three stimuli on the screen, participants had to answer by clicking on one of the two targets. 140 trials were presented.

Global

In the global version, participants had to match the global letters (and ignore the smaller one) with one of the two letters presented at the bottom of the screen by selecting one of them. Global letters were E or M and local letters were H, T or A (Fig. 9B). After the appearance of the three stimuli on the screen, participants had to answer by clicking on one of the two targets. 140 trials were presented.

3.3.3.2. Results. A 2 × 2 ANOVA was performed with Group (amusics, controls) as a between-participants factor and Condition (Local, Global) as a within-participant factor. The ANOVA revealed no significant main effect of Group [$F(1,34) = 0.12$, $p = 0.73$, $\eta_p^2 = 0.004$, Cohen's $f = 0$]. The main effect of Condition was significant [$F(1,34) = 14.99$, $p < 0.001$, $\eta_p^2 = 0.31$, Cohen's $f = 0.62$]. Participants had shorter reaction times for the Global condition than for the Local condition. No significant interaction was observed [$F(1,34) = 0.08$, $p = 0.78$, $\eta_p^2 = 0.002$, Cohen's $f = 0$] (Fig. 9C). Data from one amusic participsant was excluded from analyses due to a technical error, performance for the Global condition was not saved.

Overall, for the three visuo-spatial tasks, no differences in reaction time were observed between amusic and control individuals. As expected, we found the same pattern of performance across tests with shorter reaction times when participants had to focus their attention on the global symbol (compared to the local one).

3.4. Principal component analysis (PCA)

The PCA allowed for the joint analysis of the following 33 variables: six MBEA sub-scales (number of correct responses for each subscale), PDT (in semitones), PCD, DCI, STM (percentage of correct responses averaged across all trial types for each of these three pitch tasks), AS (total time spent in one-stream percept), EMO (percentage of correct recognitions), six Audimots scores (percentage of correct responses for each of the two noise types and each of the three difficulty levels), three 15iSSQ sub-scales, three EEAS sub-scales, weighted speed index from DeltaText, memory span from "la Baleine Paresseuse", reaction times for the two Arrow tests, the three Triplet tests and the two SIGL tests. The five missing values (see results of Audimots, copying test "la Baleine Paresseuse", Switchipido, SIGL) were replaced by the mean of the corresponding group. No clear elbow was observed on the scree plot of explained variances (Fig. 10A). It might be situated between the third and fourth dimension, but we analyzed the first ten dimensions as these ten dimensions were associated with interpretable loadings of the initial variables. The eigenvalues for these ten dimensions were above 0.99 and these ten dimensions explained 80% of the variance in our data. Based on the variables with the highest loadings (Fig. 10B), the first dimension (explaining 21.5% of the variance) corresponded to pitch and music perception (six MBEA sub-scales, PDT, PCD, DCI, STM, 15iSSQ qualities of hearing, all $r > 0.55$). The second dimension corresponded to listening effort and attention tests (15iSSQ speech, the three EEAS sub-scales, Arrow alternance, the three Triplet tasks, and the two SIGL tasks, all $r > 0.45$; PCD, DCI, and PDT were $r > 0.45$ but the loadings were higher in the first dimension, all $r > 0.55$). The third dimension corresponded to stream segregation, and pitch in speech perception (AS, EMO, Audimots CocktailParty easy, and 15iSSQ speech, all $r > 0.45$; number of copied characters, and SIGL local were $r > 0.48$ but the loadings were higher in the first dimension, all $r > 0.52$) The fourth dimension corresponded to speech perception and spatial hearing (Audimots SpeechNoise easy and difficult, and Audimots CocktailParty easy, 15iSSQ spatial hearing, all $r > 0.41$). The fifth dimension was related to spatial and visual attention (15iSSQ spatial, number of copied characters, and arrow global, all $r > 0.50$). The sixth dimension was related to speech perception in noise (Audimots CocktailParty difficult, and very difficult, all $r > 0.45$). The seventh dimension corresponds to reading abilities and copying (weighted speech index, and number of back-and-forth eye-movements, all $r > 0.52$). The eighth dimension was related to stream segregation (AS, $r = 0.48$). The ninth dimension was related to hearing-in-noise (Audimots CocktailParty difficult, $r = 0.44$). Finally, the tenth dimension was also related to hearing-in-noise

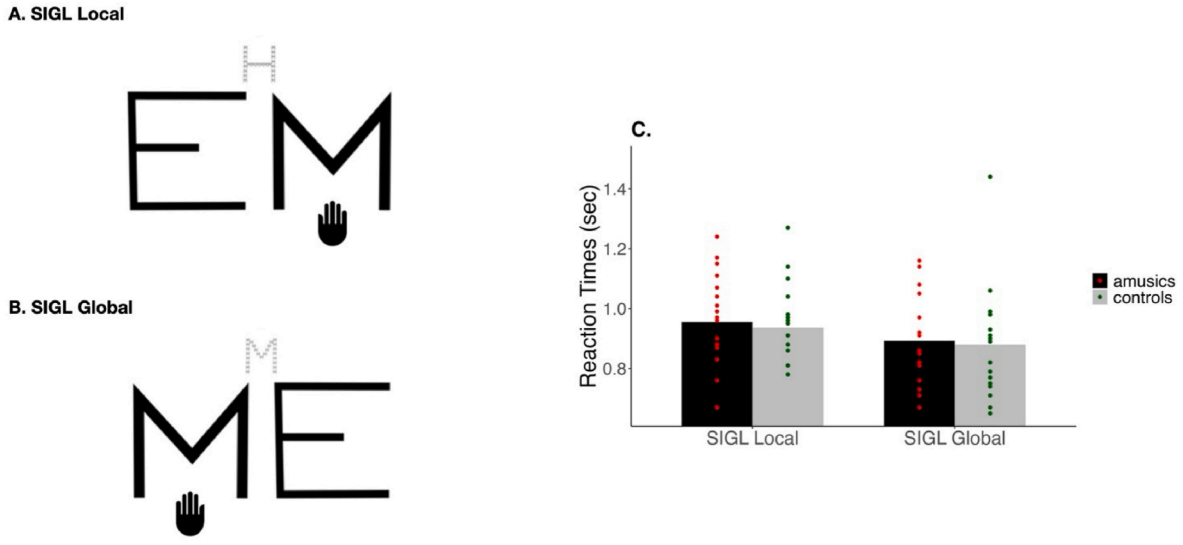


Fig. 9. SIGL tasks. **A.** In SIGL Local, the target stimulus was a hierarchical letter: a global letter made of smaller (local) letters (different from the global one), presented at the top of the screen. Participants have to match the local letters (and ignore the global one) with one of the two letters presented at the bottom of the screen by selecting one of them. **B.** In SIGL Global, participants have to match the global letter (and ignore the smaller ones) with one of the two letters presented at the bottom of the screen by selecting one of them. **C.** Reaction times (in seconds) for the two SIGL tasks. Average performance for the amusic group is indicated as a black bar and for the control group as a gray bar; dots represent individual data (amusics in red and controls in green). No significant between-group differences were observed.

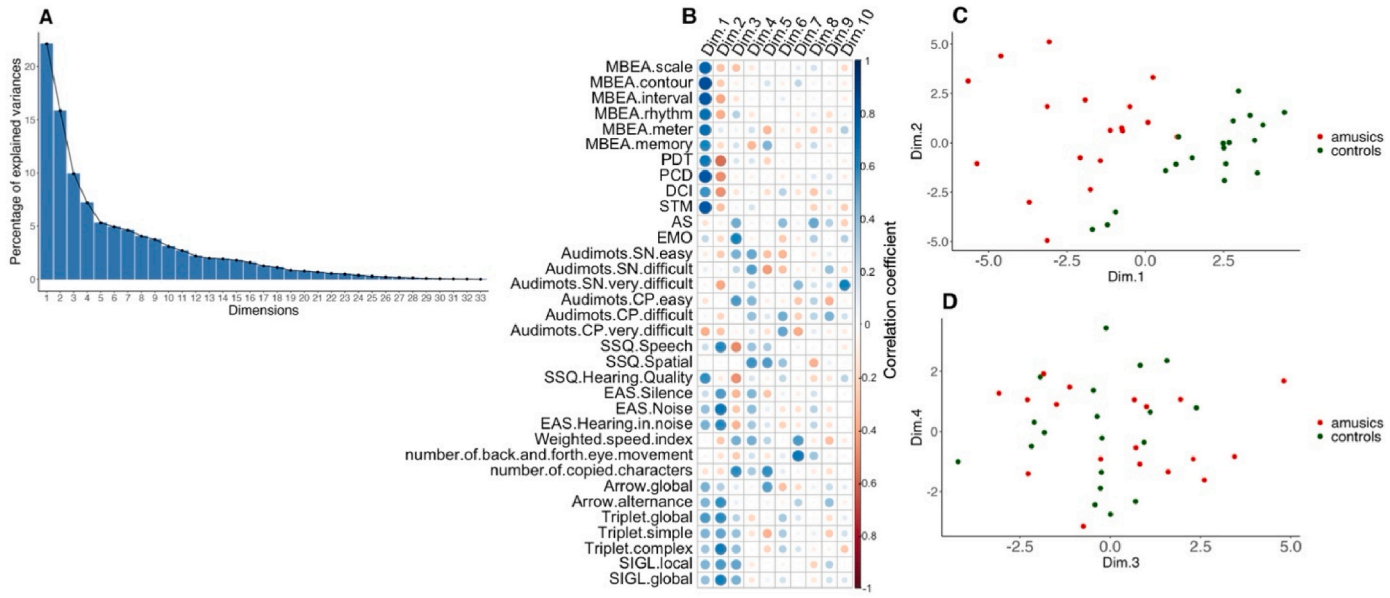


Fig. 10. Results of the Principal Component Analysis (PCA). Ten dimensions explained 80% of the variance in our data. **A.** Percentage of explained variance across dimensions. **B.** Correlation of the 33 variables with the first 10 dimensions retrieved by the PCA. The sign of the scores for PDT, AS, the three EAS sub-scales, and reaction times for the two Arrow tasks, the three Triplet tasks, and the two SIGL tasks were reversed so that higher values also correspond to better performance. **C.** Position of participants in dimensions 1 and 2 (amusic participants in red, control participants in green). Based on the loading of the variables, the first dimension corresponded to pitch and music perception (six MBEA sub-scales, PDT, PCD, DCI, STM, SSQ15 hearing qualities) and the second dimension to listening effort and attention (SSQ15 speech, the three EAS sub-scales, Arrow alternance, the three Triplet tasks, and the two SIGL tasks). **D.** Position of participants in dimensions 3 and 4. Based on the loading of the variables, the third dimension is related to pitch in speech (EMO, SSQ15 speech, and SSQ15 hearing qualities), and the fourth dimension to speech perception in noise (Audimots SpeechNoise easy and difficult, and Audimots CocktailParty easy).

(Audimots SpeechNoise very difficult, r = 0.59) (Fig. 10B).

We then compared between groups the coordinates of the participants on the PCA dimensions with Student *t* tests. A significant main effect of Group was obtained for the first dimension of the PCA (t (36) = -6.70, $p < 0.001$, Cohen's d = −2.23, remains significant after Bonferroni correction) (Fig. 10C). No significant difference between amusic and control individuals was observed for the other 9 dimensions investigated (all $p > 0.9$).

4. Discussion

Overall, the present results confirmed a pitch processing deficit in congenital amusia, notably with impaired performance for pitch change detection, pitch contour processing, and pitch short-term memory, as well as impaired emotional prosody processing in vowels. Questionnaire data revealed some first-person knowledge of these deficits, with ratings of subjectively reduced hearing quality. In contrast, stream segregation and speech-in-noise abilities did not differ between amusic and control participants. Also, for reading and visuo-spatial attention, no deficits were observed in the amusic group. However, amusic participants exhibited an atypical eye-movement strategy when copying a text, notably with a tendency to perform more numerous visits of the target text before writing, and a reduced memory span when copying a text.

4.1. Deficits of pitch processing in congenital amusia

As expected, the present results showed a pitch processing deficit in congenital amusia with significantly poorer performance in the PCD, DCI, and STM tasks. As demonstrated with the PCD, individuals with amusia detected pitch changes smaller than 0.5 semitone more poorly than did control participants who detected almost perfectly pitch changes up to 0.25 semitone. Individuals with amusia are thus less sensitive to changes in pitch than controls, confirming previous findings (Albouy et al., 2013b; Hyde and Peretz, 2004). As in Albouy et al. (2013b), significant between-group differences emerged for changes of 0.25 and 0.125 semitone. In Hyde and Peretz (2004), between-group differences emerged also for larger pitch differences (notably 1/4 and 1/2 tones). This difference might be related to the different age ranges of the populations tested in the different studies (mean age: 57–58 years in Hyde and Peretz, 2004; 35 years in Albouy et al. (2013b); 31–32 years in the current study).

In the DCI task, participants had to go beyond the detection of a pitch change, they had to distinguish whether the second tone of a pair is higher or lower than the first, creating a rising or falling contour. The task thus required processing of pitch change direction. Results showed lower performance for amusic participants than control participants, suggesting a deficit of pitch pattern processing in congenital amusia. These results are consistent with previously reported impairments in congenital amusia, such as Foxton et al. (2004) reporting deficits in a pitch direction determination task, and Liu et al. (2010) reporting deficits in pitch change and pitch direction discrimination tasks with discrete and gliding tones. Pitch direction is also relevant when differentiating intonations at the end of sentences (e.g., between a statement and a question, the final pitch direction differs). This ability was found to be impaired in congenital amusia (Liu et al., 2010), suggesting that the observed pitch direction discrimination deficit extends to speech (see Stewart, 2011, for a review).

The current study thus adds to previous findings (review in Tillmann et al., 2015, 2023) showing pitch perception deficits in congenital amusia, encompassing both the detection of changes in pitch and the identification of the direction of pitch changes (contour) and deficits in the short-term memory (STM) task (e.g., Albouy et al., 2013a; Tillmann et al., 2009). Combining results from the DCI and STM tasks thus confirms that congenital amusia is associated with a deficit of musical contour processing, observed both with a simple two-tone task (DCI) and with longer pitch sequences (STM). In the STM task, differences in accuracy between groups were observed only for "different" (and not "identical") trials, for both small (inferior to 5 semitones) and large (superior to 5 semitones) pitch changes. The short-term memory deficit thus persists even when the change in pitch is greater than individual

discrimination thresholds (the worst pitch discrimination threshold, PDT, in the amusic group is 2.93 semitones, while the large pitch changes in the STM task were 5 semitones and above), which is in agreement with previous findings (Albouy et al., 2016; Gosselin et al., 2009; Tillmann et al., 2009). Even though for "identical" trials amusic participants reached the same level of performance as controls, they exhibited longer response times, which is consistent with other studies using different tasks and sounds (e.g., Albouy et al., 2013b; Tillmann et al., 2014). These findings suggest that individuals with amusia are more uncertain and therefore might need more time to reach a decision. Here, we further investigated the effect of sequence length (four tones vs. six tones) on memory performance: increasing sequence length impaired performance both for amusic and control participants. Gosselin et al. (2009) observed a non-significant numerical tendency for a larger effect of increasing tone sequence length (from 1 to 5 tones) on memory performance in individuals with amusia compared to control participants. Here we also did not observe a significant interaction between group and sequence length on memory performance, note however that the effect size for the main effect of the group factor was larger for the STM task with six-tone sequences ($\eta_p^2 = 0.48$, Cohen's f = 0.92) than for the STM task with four-tone sequences ($\eta_p^2 = 0.23$, Cohen's f = 0.5). Larger between-group differences in short-term memory performance and no overlap in the performance of the two groups were observed in previous studies (Albouy et al., 2013a; Tillmann et al., 2009). This absence of overlap between groups could be explained by more challenging conditions in the previous STM tasks than in the STM task used here. Indeed, STM tasks with five-tone sequences and with changes in different trials being created by exchanging tone position in the second sequence (without necessarily entailing a contour change between the melodies to compare, which was the case here) (Tillmann et al., 2009) might be more challenging than the current implementation. Furthermore, reducing the SOA between the tones of the melodic sequences also had a pronounced deleterious impact on performance, in particular for amusic participants (Albouy et al., 2016; Albouy et al., 2013a). Future studies should aim to delineate the characteristics of the stimuli that are most detrimental to short-term memory performance in congenital amusia (rapid auditory sequences, memory load, duration of the retention interval, interferents, presence of contour changes) and assess the combined impact of these different factors.

Overall, participants with congenital amusia had impaired performance in pitch-related tasks, and this deficit is not only due to a simple pitch discrimination impairment, but is also related to a pitch memory impairment (Tillmann et al., 2009, 2016). In keeping with this interpretation and as expected, we observed a large overlap in pitch discrimination thresholds of amusic and control participants (Table 1). Across all the tests performed here, we find the largest effect sizes for the main effect of group for the six-tone sequence short-term-memory task ($\eta_p^2 = 0.48$, Cohen's f = 0.92), the pitch tasks (PCD, DCI, STM; $\eta_p^2 = 0.43$, Cohen's f = 0.88) and the PCD ($\eta_p^2 = 0.51$, Cohen's f = 0.98) supporting the importance of pitch deficits, in particular in challenging memory tasks, in congenital amusia. The large effect sizes observed in our studies for the pitch tasks are consistent with previous studies which reported medium to large between-group differences (see Table 2). The deficit in pitch short-term memory could explain the perception and contour deficit in congenital amusia (Stewart, 2011; Tillmann et al., 2009, 2015, 2023). For example, individuals need to keep in memory the previous tones to determine whether the contour goes upwards or downwards.

Interestingly, the deficit of pitch perception and memory observed in congenital amusia might have a self-perceived impact on amusic participants' quality of hearing. Indeed, we observed significant group differences for the two questions on music in the qualities of Hearing sub-scale (15iSSQ15 questionnaire): "Can you easily recognize the different pieces of music you know?" and "When you listen to music, does it sound clear and natural to you?", with lower scores in individuals with amusia than the control participants. Moreover, the qualities of Hearing sub-scale was associated with pitch tasks (PCD, DCI, and STM) in the first dimension of the PCA.

4.2. Deficits of emotional prosody processing in congenital amusia

As pitch is essential in prosody, we investigated emotional prosody perception in congenital amusia. We used sentences and vowels pronounced with different emotions (Joy, Sadness, Anger, Fear, Neutral) to investigate emotion recognition and the assessment of their perceived intensity (replicating Pralus et al., 2019). Our results confirmed that congenital amusic participants had difficulty recognizing emotion in vowels, but not in sentences, in agreement with previous results using the same material (Pralus et al., 2019). Indeed, emotion recognition could be further supported by the long duration of the sentences (1470ms for sentences vs. 400ms for vowels) and by other acoustic cues, such as intensity and duration cues (see Pralus et al., 2019 for acoustic analyses).

Analyzing emotion intensity ratings allowed us to investigate subjective assessment of the emotions and possibly more implicit processes than those involved in the categorization task (Lévêque et al., 2018; Pralus et al., 2019; Pralus et al., 2020b). For emotion perception in sentences and vowels, the intensity ratings of participants with amusia did not differ from those of control participants, as in Pralus et al. (2019), and as also observed for music (Lévêque et al., 2018).

Overall, the present results obtained with the pitch tasks and emotion-related tasks were in agreement with findings from previous studies (Albouy et al., 2015b; Foxton et al., 2004; Pralus et al., 2019; Tillmann et al., 2009), and thus confirmed deficits of pitch processing for musical and verbal sounds in congenital amusia, including emotion recognition in verbal material based on pitch and other spectral cues (Lima et al., 2016; Pralus et al., 2019; Thompson et al., 2012). In addition, previous studies have investigated tone-language processing in congenital amusia and have revealed impaired performance in pitch discrimination in tone-language speech (Nan et al., 2010; Tillmann et al., 2011a) as well as spoken syllables manipulated in pitch height (Tillmann et al., 2011b), supporting the hypothesis of pitch deficits that extend to speech processing in congenital amusia.

4.3. Preserved stream segregation and hearing-in-noise in congenital amusia

As stream segregation is related to pitch processing, we investigated its potential impairment in congenital amusia. The AS test allows for investigating the ability to separate sound sources (Grimault et al., 2002). No difference was observed between amusic and control participants regarding the time spent in one stream and two streams percepts. Participants in both groups spent most of the time in the two streams percept, as did control participants in Pralus et al. (2021). No difference was observed between amusic and control participants for the time spent in each percept and the mean frequency at the change of percept (a low mean frequency of change suggests good segregation ability), suggesting that the way pitch is organized and used for stream segregation is not impaired in congenital amusia. These results are consistent with the study of Foxton et al. (2004) who detected no abnormalities in the way pitch is perceptually organized using a similar task. Our present set of tasks allowed us to further test, in the same participants, whether preserved stream segregation abilities are associated with intact performance in a speech-in-noise test and self-perceived difficulties in noisy conditions assessed with questionnaires.

As separation of sound sources, notably based on pitch cues, is essential for hearing in a noisy environment, we investigated speech-in-noise abilities in congenital amusia using the Audimots test (Moulin et al., 2013) with two types of noise and three levels of phonological difficulty. As expected, our results revealed a main effect of noise, with higher performance with SpeechNoise compared to CocktailParty, the former entailing mostly energetic masking and the latter both energetic and informational masking. In addition, the phonological similarity

between the words significantly influenced participants' performance. Participants of both groups performed worse when the words had similar phonetics, particularly in trials focusing on initial consonant distinctions compared to trials involving only vowel differences. We observed in the two noise conditions (SpeechNoise and CocktailParty) a floor effect for trials focusing on initial consonant differences for both amusic and control participants. Indeed, words differing by only the initial consonants are more difficult to distinguish, as consonants are both shorter and of lower amplitude than vowels. Also, for word perception in noise, we observed no difference between amusic and control participants. This finding is in line with the studies of W. Tang et al. (2018) and Loutrari et al. (2024) who also did not find deficits for tone or speech perception in noise. Except Liu et al. (2015a) and Shao et al. (2016) who reported a group difference in speech-in-noise performance with a large effect size (Cohen's $f > 0.68$), previous studies and the present one revealed only non-significant group differences and medium to small effect sizes (Cohen's $f < 0.5$, see Table 2). It seems that detecting between-group differences in speech-in-noise processing, if any, with a power of 0.7, would require rather larger sample sizes than usually tested in congenital amusia studies ($n > 350$ to detect reliably small between-group differences). Our results thus suggest that pitch perception deficit in amusia does not seem to affect (at least not to a large extent) their ability to analyze auditory scenes in terms of different streams, including to extract a target stimulus from noise. However, our results are divergent from those of Liu et al. (2015a). In this previous study, participants were tested with sentences presented in quiet and in noise with different SNR condition, and were asked to write the words they heard on a paper. Additional studies will be necessary to investigate these differences, particularly by exploring potential subgroups among individuals with amusia (with or without speech tone difficulties as in W. Tang et al. (2018) or by investigating different materials (vowels, sentences, type of noises, SNR) and procedures (repetition/transcription vs. forced-choice methods, such as used here, which rely less on memory and production abilities).

These results in listening tests were consistent with the scores obtained at the 15iSSQ speech perception sub-scale and the EEAS questionnaire, which both did not reveal any group differences: The 15iSSQ speech perception sub-scale assesses speech perception in noisy environments and the EEAS questionnaire measures the self-reported listening effort in noisy and quiet environments. Participants with amusia reported the same listening effort as did control participants to each sub-scale (hearing in silence, noise, and hearing in noise). Both amusic and control participants reported more listening effort for items related to hearing in noise than for items related to hearing in quiet sub-scales, a finding in agreement with Ferschneider and Moulin (2023).

Overall, these results (AS test, Audimots, questionnaires) did not reveal any impairment of stream segregation and hearing-in-noise abilities in congenital amusia, suggesting at least partially distinct mechanisms between stream segregation on the one hand and pitch processing in music and speech on the other hand. A possible interpretation of this finding is that stream segregation, including speech-in-noise perception, is supported in part by subcortical mechanisms (Coffey et al., 2017). Indeed, Liu et al., 2015b studied subcortical encoding of speech in congenital amusia by investigating frequency-following responses (FFRs) in a passive listening task in noise and observed no group differences on FFR amplitude, suggesting normal sound processing at the subcortical level in congenital amusia. However, in older individuals with amusia (>60yo), a reduced auditory brainstem response to speech sounds is observed compared to control participants (Lehmann et al., 2015). In future studies, it could be interesting to investigate stream segregation with a more fine-grained approach, using both pitch and timbre cues, and controlling for participants' age and peripheral hearing status. Indeed, timbre is also important in stream segregation (Bregman, 1994) and timbre processing is altered in congenital amusia (Graves et al., 2019; Marin et al., 2012; Tillmann et al., 2009). Depending on the context, control participants could benefit from timbre cues compared to participants with amusia (see also Tillmann et al., 2009).

Thus, by investigating pitch in several domains, the current study shed more light on auditory perception and cognition, more specifically on the links between these different abilities relying on pitch, suggesting some degree of dissociation between stream segregation and other aspects of pitch-related auditory cognition.

4.4. On the relationship between music and language disorders

In the present study, we tested amusic individuals without dyslexia. Based on Couvignou and collaborators' reports on potential enhanced comorbidity between amusia and dyslexia (Couvignou et al., 2019, 2023; Couvignou and Kolinsky, 2021), we set out to include a test that investigated text reading and copying abilities, aiming to test whether amusia might also affect these performances even without reaching a diagnosis of dyslexia. The reading test did not reveal any difference between amusic and control participants for reading time, number of reading errors and weighted speed index. The absence of reading impairment in the current sample of amusic participants is in keeping with the fact that we only recruited participants that self-reported to not have a diagnosis of dyslexia. This was indeed an exclusion criterion for our aim to reveal amusia-specific deficits without dyslexia-related deficits. For the copying test, the numbers of characters correctly copied did not differ between participants with amusia and control participants. However, amusic participants did more back-and-forth eye-movements than did controls (Cohen's $d = 0.75$), corresponding to a smaller memory span (Cohen's $d = -0.7$). While none of the amusic participants reported having reading or writing impairments, they used a different strategy to copy the same number of characters than did controls in a given time, which could reflect a lower memory span, or a need for more verifications (lower confidence in their performance). Previous studies found no reduced memory span in congenital amusia using digit span tests (Albouy et al., 2013b; Peretz et al., 2002; Williamson and Stewart, 2010). However, the digit span task uses oral restitution and participants are asked explicitly to memorize the items which is not the case in our copying task. This could explain the reduced memory span observed with the more naturalistic copying test and not with the oral digit span test. However, we should note that even if more back-and-forth eye--movements and lower memory span are observed in individuals with amusia, effect sizes were smaller (medium effect size) than for pitch tasks (large effect sizes). In addition, reading and copying measures come out on a separate dimension from pitch tasks in the PCA analysis, suggesting possible subgroups in the amusic participants.

The behavior of participants with amusia in the copying test might be related to a sequential memory impairment rather than impaired visuo-spatial attention. Investigating visuo-spatial attention in amusia is interesting as dyslexic individuals may present visuo-attentional deficits (Bedoin et al., 2010; Brannan and Williams, 1987). We used visual control tasks measuring the ability to switch between local and global attention. As expected, participant groups did not differ for any of the used measures, that is reaction time for the Switchipido (Arrow, Triplet) and SIGL tasks, confirming no visuo-attentional deficits in congenital amusia. In both amusic and control participants we observed the classical performance pattern in these tasks. Indeed, for the three tasks, Arrow, Triplet, and SIGL, participants presented shorter reaction times when they had to focus their attention on the global symbol, reflecting automatic global symbol processing. When participants had to focus their attention on local symbols, the presence of a global symbol creates an interference. These results are in line with the literature (Bedoin et al., 2010) showing that in typically-developing adults, interference decreases performance and is more pronounced when it comes from the global level than from the local level reflecting the "interference asymmetry". Moreover, these results are in line with previous studies revealing no spatial processing deficits in congenital amusia (Tillmann et al., 2010; Williamson et al., 2011).

Our findings resonate with the work of Couvignou et al. (2023) on the comorbidity between amusia and dyslexia. As in Couvignou et al. (2023), we highlight memory processes that would be shared between music and language, possibly in relation with time-based serial-order processing, that is critical for both music and language, here reflected in the performance in the STM task for music and the copying task for language.

Congenital amusia and dyslexia affect primarily different facets of information processing (musical vs. linguistic), but seem to share some similarities in terms of cognitive functioning. The mechanisms underlying the comorbidity between amusia and dyslexia are complex and not fully understood yet, thus connection between these two disorders still needs further research. To go further, reading behavior in individuals with amusia could be tested with more fine-grained methods, such as eye-tracking in order to analyze the sequential behavior in amusia and link this sequential behavior to memory abilities (in particular, serial order short-term memory). In addition, it could be interesting to investigate profiles of deficits in individuals with congenital amusia presenting a comorbidity with dyslexia, especially with tasks involving verbal and pitch memory, to provide new insights on commonly impaired cognitive mechanisms.

4.5. Patterns of deficits in peripheral and central auditory disorders

In the present study, thanks to the variety of tests used, we provided new insights on the pattern of deficits in congenital amusia, and compared it to other neurodevelopmental disorders such as dyslexia. We can further compare this profile to the pitch processing impairments resulting from peripheral auditory deficits (hearing impairment without lesions) as well as central pitch processing deficits in particular in patients with brain lesions.

We analyzed the pattern of deficits in congenital amusia across our tasks with the Principal Component Analysis. The first dimension recovered correlates with the six MBEA sub-scales, the pitch detection threshold, pitch change detection, pitch direction of identification, short-term memory, and 15iSSQ hearing qualities sub-scale. Individuals with amusia were impaired for all these tasks compared to control participants and the two groups are quite well separated on the first PCA dimension (Fig. 10C). The other dimensions retrieved by the PCA reflect the other cognitive functions under scrutiny here (Fig. 10B): auditory scene analysis (including speech-in-noise), attention, reading and copying. None of these other dimensions allowed separating the two groups. The underlying deficit in congenital amusia thus appears as quite homogeneous across participants and tasks and is related to pitch processing only. Note, however, that our investigation of temporal processing and rhythm was limited (two MBEA subtests, Rhythm and Meter, explore them, but in melodically rich contexts). These time-based processes should be more investigated, notably when studying comorbidity between amusia and dyslexia, given the link between short-term memory of serial order on the one hand, and temporal and rhythm processing on the other hand (Gorin et al., 2016).

Overall, amusic individuals exhibit deficits in pitch perception and short-term memory tasks (Albouy et al., 2013a; Tillmann et al., 2016, 2023), emotion recognition in prosody and music (Lévêque et al., 2018; Pralus et al., 2019) revealing impaired processing of pitch and spectral content, but limited if any deficits in auditory scene analysis (stream segregation, speech-in-noise perception). Congenital amusia is characterized by deficits at the cortical level (Albouy et al., 2015b; Albouy et al., 2013a; Hyde et al., 2006, 2007; Leveque et al., 2016; Moreau et al., 2013; Peretz et al., 2005) with relatively preserved sub-cortical processing (Lehmann et al., 2015; Liu et al., 2015b) and no audiometric hearing loss.

This pattern of deficits for congenital amusia can be compared to deficits in clinical populations, notably hearing-impaired participants wearing cochlear implant(s) (CI) and brain-damaged patients, both tested with subsets of the tests used here (Pralus et al., 2021 for CI users; Hirel et al., 2017; Pralus et al., 2020a, for brain-damaged patients). In CI users, pitch processing is altered because of the limited frequency resolution of the implant. CI users are impaired in pitch change detection, recognition of emotional prosody in sentences, stream segregation, and hearing-in-noise abilities (Hong and Turner, 2006; Oxenham, 2008; Pralus et al., 2021). The comparison between individuals with congenital amusia and CI users opposes central deficits of pitch perception to peripheral deficits, with overall more pronounced deficits in stream segregation and hearing-in-noise in CI users, differences that might reflect both the degraded peripheral input, which will impact heavily also on sound representation in subcortical structures, and possible compensations at a more central level of processing in CI users.

Regarding acquired musical deficits, left-brain damaged patients present musical emotion recognition (Pralus et al., 2020a) and pitch short-term memory deficits (Hirel et al., 2017), with preserved emotion intensity judgments (as in congenital amusia, Lévêque et al., 2018, and in Landau-Kleffner epileptic syndrome, Lévêque et al., 2020). Right brain damage patients exhibit no deficit in emotion recognition but are impaired in emotion intensity judgments (Pralus et al., 2020a) and in pitch short-term memory (Hirel et al., 2017). Thus, across populations with neurodevelopmental (congenital amusia, congenital musical anhedonia) and acquired (brain lesions) central auditory processing deficits, we observe a notable dissociation between explicit emotion labeling and perceived intensity of emotions (Mas-Herrero et al., 2014).

These studies on auditory cognition through different auditory disorders highlight the importance of pitch processing in music perception and emotion recognition, even for (short) verbal materials. These processes can be altered in different conditions, with various profiles of impairments.

4.6. Limitations of the present study

All participants reported no diagnosis of neurological or psychiatric disorders and had all followed classic schooling leading to the A-level. However, we did not include classical neuropsychological tests (such as a Digit span task) in the tests administered, even though it would have been interesting to have an assessment of other cognitive abilities. Note that in a previous study from our group, which was testing participants with similar backgrounds, performance on the digit span task did not differ between the groups (Albouy et al., 2013b). Statistical power could be a limitation due to the number of tasks used and the size of the participant samples. We thus performed statistical power analyses based on the effect sizes observed in previous studies using the same or similar tasks (see Table 2) to analyze this limit. Given the current sample sizes, we had good power to detect large between-group differences. Detecting smaller to medium effects would require such large cohorts that multi-lab studies or meta-analyses would probably be necessary given the prevalence of congenital amusia. Another limitation in our study is the counting of back-and-forth eye-movements during the copying test. Counting was done only by one experimenter (not blinded the participant's group) and the experiment was not videotaped, thus not allowing for a double coding by another experimenter. Future studies should rather record eye-movements (e.g., using eye-tracking) to obtain more objective measures, which also allow for blind coding.

4.7. Perspectives

The present study allowed us to further characterize the deficits observed in congenital amusia and to compare these deficits with those observed in other neurodevelopmental disorders, such as dyslexia, or other auditory disorders observed in patients with brain damage or cochlear implant users.

As some of the pitch-related deficits are shared among congenital

amusia and other auditory disorders, developing remediation programs targeting pitch perception and memory could be valuable in amusia and beyond (see Whiteford and Oxenham, 2018 for an improvement of pitch processing in individuals with amusia after an auditory training protocol targeting pure-tone pitch discrimination).

CRediT authorship contribution statement

Caliani Hoarau: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Agathe Pralus:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Annie Moulin:** Writing – review & editing, Software, Resources, Methodology, Conceptualization. **Nathalie Bedoin:** Writing – review & editing, Resources, Methodology, Conceptualization. **Jérémie Ginzburg:** Writing – review & editing, Supervision, Investigation. **Lesly Fornoni:** Writing – review & editing, Supervision, Project administration, Investigation, Data curation. **Pierre-Emmanuel Aguera:** Writing – review & editing, Software. **Barbara Tillmann:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Anne Caclin:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

CH is funded by a CIFRE PhD fellowship from Humans Matter and ANRT. Humans Matter holds an exclusive license to commercialize the three pitch tests used here (PCD, DCI, MCT) but had no role in study design and data analysis.

Data availability

Data will be made available on request.

Acknowledgements

This work was funded by a grant of the Fédération pour la Recherche sur le Cerveau (FRC) to AC, AM, NB, and BT. This work was conducted within the framework of the LabEx CeLyA (“Centre Lyonnais d’Acoustique”, ANR-10-LABX-0060) of Université de Lyon, within the program “Investissements d’avenir” (ANR-16-IDEX-0005) operated by the French National Research Agency (ANR). We thank Antoine Dehem for his help with data collection and Jackson Graves for proofreading our manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2024.108960.

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. Sci. Rep. 6 (1), 18861 https://doi.org/10.1038/srep18861.

Albouy, P., Lévêque, Y., Hyde, K.L., Bouchet, P., Tillmann, B., Caclin, A., 2015a. Boosting pitch encoding with audiovisual interactions in congenital amusia. Neuropsychologia 67, 111–120. https://doi.org/10.1016/j.neuropsychologia.2014.12.006.

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P.-E., Daligault, S., Delpuech, C., Bertrand, O., Caclin, A., Tillmann, B., 2013a. 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., 2015b. Altered retrieval of melodic information in congenital amusia: insights from dynamic causal modeling of MEG data. Front. Hum. Neurosci. 9 https://doi.org/10.3389/fnhum.2015.00020.

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

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

Alhanbali, S., Dawes, P., Lloyd, S., Munro, K.J., 2017. Self-reported listening-related effort and fatigue in hearing-impaired adults. Ear Hear. 38 (1), e39–e48. https://doi.org/10.1097/AUD.0000000000000361.

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

Bedoin, N., 2017. Rééquilibrer les analyses visuo, vol. 10. Attentionnelles globales et locales pour améliorer la lecture chez des enfants dyslexiques de surface.

Bedoin, N., Kéïta, L., Leculier, L., Roussel, C., Herbillon, V., Launay, L., 2010. Diagnostic et remédiation d’un déficit d’inhibition des détails dans la dyslexie de surface. In: Rousseau, T., Valette-Fruhinsholz, F. (Eds.), Le Langage Oral: Données Actuelles et Perspectives En Orthophonie. *Isbergues: OrthoEditions*.

Bedoin, N., Medina, F., 2013. Switchipido, training software for alternating between local and global visual treatments, and for the inhibition of details [Computer software]. www.gnosia.fr.

Boets, B., Op De Beeck, H.P., Vandermosten, M., Scott, S.K., Gillebert, C.R., Mantini, D., Bulthé, J., Sunaert, S., Wouters, J., Ghesquière, P., 2013. Intact but less accessible phonetic representations in adults with dyslexia. Science 342 (6163), 1251–1254. https://doi.org/10.1126/science.1244333.

Bourgeois–Vionnet, J., Moulin, A., Hermier, M., Pralus, A., Nighoghossian, N., 2020. A case of verbal and emotional prosody processing dissociation after a right temporal venous infarct. Neurol. Sci. 41 (6), 1615–1618. https://doi.org/10.1007/s10072-020-04259-y.

Brannan, J.R., Williams, M.C., 1987. Allocation of visual attention in good and poor readers. Percept. Psychophys. 41 (1), 23–28. https://doi.org/10.3758/BF03208209.

Bregman, A.S., 1994. Auditory Scene Analysis. MIT Press, Cambridge, MA, p. 198.

Charpentier, J., Kovarski, K., Roux, S., Houy-Durand, E., Saby, A., Bonnet-Brilhault, F., Latinus, M., Gomot, M., 2018. Brain mechanisms involved in angry prosody change detection in school-age children and adults, revealed by electrophysiology. Cognit. Affect Behav. Neurosci. 18 (4), 748–763. https://doi.org/10.3758/s13415-018-0602-8.

Coffey, E.B.J., Chepesiuk, A.M.P., Herholz, S.C., Baillet, S., Zatorre, R.J., 2017. Neural correlates of early sound encoding and their relationship to speech-in-noise perception. Front. Neurosci. 11, 479. https://doi.org/10.3389/fnins.2017.00479.

Corrow, S.L., Stubbs, J.L., Schlaug, G., Buss, S., Paquette, S., Duchaine, B., Barton, J.J.S., 2019. Perception of musical pitch in developmental prosopagnosia. Neuropsychologia 124, 87–97. https://doi.org/10.1016/j.neuropsychologia.2018.12.022.

Couvignou, M., Kolinsky, R., 2021. Comorbidity and cognitive overlap between developmental dyslexia and congenital amusia in children. Neuropsychologia 155, 107811. https://doi.org/10.1016/j.neuropsychologia.2021.107811.

Couvignou, M., Peretz, I., Ramus, F., 2019. Comorbidity and cognitive overlap between developmental dyslexia and congenital amusia. Cogn. Neuropsychol. 36 (1–2), 1–17. https://doi.org/10.1080/02643294.2019.1578205.

Couvignou, M., Tillmann, B., Caclin, A., Kolinsky, R., 2023. Do developmental dyslexia and congenital amusia share underlying impairments? Child Neuropsychol. 1–47. https://doi.org/10.1080/09297049.2022.2162031.

Douglas, K.M., Bilkey, D.K., 2007. Amusia is associated with deficits in spatial processing. Nat. Neurosci. 10 (7), 915–921. https://doi.org/10.1038/nn1925.

Duplat, A., Girier, J., 2006. Étude de la correlation entre fenêtre attentionnelle et fenêtre de copie chez des enfants normolecteurs et dyslexiques. Mémoire D’orthophonie.

Ferschneider, M., Moulin, A., 2023. Listening effort in quiet and noisy environments in the daily life of adults with hearing aids: an extended version of the effort assessment scale (EEAS). Trends in Hearing 27, 23312165231176320. https://doi.org/10.1177/23312165231176320.

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.

Gatehouse, S., Noble, W., 2004. The speech, spatial and qualities of hearing scale (SSQ). Int. J. Audiol. 43 (2), 85–99. https://doi.org/10.1080/14992020400050014.

Gorin, S., Kowialiewski, B., Majerus, S., 2016. Domain-generality of timing-based serial order processes in short-term memory: new insights from musical and verbal domains. PLoS One 11 (12), e0168699. https://doi.org/10.1371/journal.pone.0168699.

Gosselin, N., Jolicœur, P., Peretz, I., 2009. Impaired memory for pitch in congenital amusia. Ann. N. Y. Acad. Sci. 1169 (1), 270–272. https://doi.org/10.1111/j.1749-6632.2009.04762.x.

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

Grimault, N., Bacon, S.P., Micheyl, C., 2002. Auditory stream segregation on the basis of amplitude- modulation ratea). J. Acoust. Soc. Am. 111 (3).

Hirel, C., Nighoghossian, N., Lévêque, Y., Hannoun, S., Fornoni, L., Daligault, S., Bouchet, P., Jung, J., Tillmann, B., Caclin, A., 2017. Verbal and musical short-term

memory: variety of auditory disorders after stroke. Brain Cognit. 113, 10–22. https://doi.org/10.1016/j.bandc.2017.01.003.
Hong, R.S., Turner, C.W., 2006. Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. J. Acoust. Soc. Am. 120 (1), 360–374. https://doi.org/10.1121/1.2204450.
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. J. Neurosci. 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. Psychol. Sci. 15 (5), 356–360. https://doi.org/10.1111/j.0956-7976.2004.00683.x.
Hyde, K.L., Zatorre, R.J., Griffiths, T.D., Lerch, J.P., Peretz, I., 2006. Morphometry of the amusic brain: a two-site study. Brain 129 (10), 2562–2570. https://doi.org/10.1093/brain/awl204.
Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an *R* package for multivariate analysis. J. Stat. Software 25 (1). https://doi.org/10.18637/jss.v025.i01.
Lebrun, M.-A., Moreau, P., McNally-Gagnon, A., Mignault Goulet, G., Peretz, I., 2012. Congenital amusia in childhood: a case study. Cortex 48 (6), 683–688. https://doi.org/10.1016/j.cortex.2011.02.018.
Lehmann, A., Skoe, E., Moreau, P., Peretz, I., Kraus, N., 2015. Impairments in musical abilities reflected in the auditory brainstem: evidence from congenital amusia. Eur. J. Neurosci. 42 (1), 1644–1650. https://doi.org/10.1111/ejn.12931.
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. J. Neurophysiol. 116 (1), 88–97. https://doi.org/10.1152/jn.00663.2015.
Lévêque, Y., Roulet-Perez, E., Deonna, T., Moulin, A., Fornoni, L., Mayor-Dubois, C., Caclin, A., Tillmann, B., 2020. Music processing deficits in Landau-Kleffner syndrome: four case studies in adulthood. Cortex 129, 99–111. https://doi.org/10.1016/j.cortex.2020.03.025.
Lévêque, Y., Teyssier, P., Bouchet, P., Bigand, E., Caclin, A., Tillmann, B., 2018. Musical emotions in congenital amusia: impaired recognition, but preserved emotional intensity. Neuropsychology 32 (7), 880–894. https://doi.org/10.1037/neu0000461.
Lima, C.F., Brancatisano, O., Fancourt, A., Müllensiefen, D., Scott, S.K., Warren, J.D., Stewart, L., 2016. Impaired socio-emotional processing in a developmental music disorder. Sci. Rep. 6 (1), 34911 https://doi.org/10.1038/srep34911.
Liu, F., Jiang, C., Francart, T., Chan, A.H.D., Wong, P.C.M., 2017. Perceptual learning of pitch direction in congenital amusia. Music Perception 34 (3), 335–351. https://doi.org/10.1525/mp.2017.34.3.335.
Liu, F., Jiang, C., Wang, B., Xu, Y., Patel, A.D., 2015a. A music perception disorder (congenital amusia) influences speech comprehension. Neuropsychologia 66, 111–118. https://doi.org/10.1016/j.neuropsychologia.2014.11.001.
Liu, F., Maggu, A.R., Lau, J.C.Y., Wong, P.C.M., 2015b. Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers. Front. Hum. Neurosci. 8 https://doi.org/10.3389/fnhum.2014.01029.
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.
Liu, F., Xu, Y., Patel, A.D., Francart, T., Jiang, C., 2012. Differential recognition of pitch patterns in discrete and gliding stimuli in congenital amusia: evidence from Mandarin speakers. Brain Cognit. 79 (3), 209–215. https://doi.org/10.1016/j.bandc.2012.03.008.
Lolli, S.L., Lewenstein, A.D., Basurto, J., Winnik, S., Loui, P., 2015. Sound frequency affects speech emotion perception: results from congenital amusia. Front. Psychol. 6 https://doi.org/10.3389/fpsyg.2015.01340.
Loui, P., Guenther, F.H., Mathys, C., Schlaug, G., 2008. Action–perception mismatch in tone-deafness. Curr. Biol. 18 (8), R331–R332. https://doi.org/10.1016/j.cub.2008.02.045.
Loutrari, A., Alqadi, A., Jiang, C., Liu, F., 2024. Exploring the role of singing, semantics, and amusia screening in speech-in-noise perception in musicians and non-musicians. Cognit. Process. 25 (1), 147–161. https://doi.org/10.1007/s10339-023-01165-x.
Lu, X., Ho, H.T., Liu, F., Wu, D., Thompson, W.F., 2015. Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: an ERP study. Front. Psychol. 6 https://doi.org/10.3389/fpsyg.2015.00385.
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.
Mas-Herrero, E., Zatorre, R.J., Rodriguez-Fornells, A., Marco-Pallarés, J., 2014. Dissociation between musical and monetary reward responses in specific musical anhedonia. Curr. Biol. 24 (6), 699–704. https://doi.org/10.1016/j.cub.2014.01.068.
Mignault Goulet, G., Moreau, P., Robitaille, N., Peretz, I., 2012. Congenital amusia persists in the developing brain after daily music listening. PLoS ONE 7 (5), e36860. https://doi.org/10.1371/journal.pone.0036860.
Moreau, P., Jolicœur, P., Peretz, I., 2013. Pitch discrimination without awareness in congenital amusia: evidence from event-related potentials. Brain Cognit. 81 (3), 337–344. https://doi.org/10.1016/j.bandc.2013.01.004.
Moulin, A., Garcia, S., Jeanvoine, A., Richard, C., 2013. Audimots": Development of a French version of the Four Alternative Auditory Features Test. 11th European Federation of Audiology Societies (EFAS) Congress. Otorhinolaryngologia Hungarica 59 (2), 98–99.
Moulin, A., Pauzie, A., Richard, C., 2015. Validation of a French translation of the speech, spatial, and qualities of hearing scale (SSQ) and comparison with other language versions. Int. J. Audiol. 54 (12), 889–898. https://doi.org/10.3109/14992027.2015.1054040.
Moulin, A., Vergne, J., Gallego, S., Micheyl, C., 2019. A new speech, spatial, and qualities of hearing scale short-form: factor, cluster, and comparative analyses. Ear Hear. 40 (4), 938–950. https://doi.org/10.1097/AUD.0000000000000675.
Nan, Y., Sun, Y., Peretz, I., 2010. Congenital amusia in speakers of a tone language: association with lexical tone agnosia. Brain 133 (9), 2635–2642. https://doi.org/10.1093/brain/awq178.
Nooteboom, S. (n.d.). The prosody of speech: melody and rhythm. The Handbook of Phonetic Sciences, 5, 640–673.
Oxenham, A.J., 2008. Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. Trends Amplif. 12 (4), 316–331. https://doi.org/10.1177/1084713808325881.
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 (4), 357–368. https://doi.org/10.1525/mp.2008.25.4.357.
Peretz, I., 2013. The biological foundations of music. In: The Psychology of Music. Elsevier, pp. 551–564. https://doi.org/10.1016/B978-0-12-381460-9.00013-4.
Peretz, I., 2016. Neurobiology of congenital amusia. Trends Cognit. Sci. 20 (11), 857–867. https://doi.org/10.1016/j.tics.2016.09.002.
Peretz, I., Ayotte, J., Zatorre, R.J., Mehler, J., Ahad, P., Penhune, V.B., Jutras, B., 2002. Congenital amusia. Neuron 33 (2), 185–191. https://doi.org/10.1016/S0896-6273(01)00580-3.
Peretz, I., Brattico, E., Tervaniemi, M., 2005. Abnormal electrical brain responses to pitch in congenital amusia. Ann. Neurol. 58 (3), 478–482. https://doi.org/10.1002/ana.20606.
Peretz, I., Champod, A.S., Hyde, K., 2003. Varieties of musical disorders. Ann. N. Y. Acad. Sci. 999 (1), 58–75. https://doi.org/10.1196/annals.1284.006.
Peretz, I., Hyde, K.L., 2003. What is specific to music processing? Insights from congenital amusia. Trends Cognit. Sci. 7 (8), 362–367. https://doi.org/10.1016/S1364-6613(03)00150-5.
Peretz, I., Vuvan, D.T., 2017. Prevalence of congenital amusia. Eur. J. Hum. Genet. 25 (5), 625–630. https://doi.org/10.1038/ejhg.2017.15.
Pihan, H., 2006. Affective and linguistic processing of speech prosody: DC potential studies. In: Progress in Brain Research, vol. 156. Elsevier, pp. 269–284. https://doi.org/10.1016/S0079-6123(06)56014-5.
Pralus, A., Belfi, A., Hirel, C., Lévêque, Y., Fornoni, L., Bigand, E., Jung, J., Tranel, D., Nighoghossian, N., Tillmann, B., Caclin, A., 2020a. Recognition of musical emotions and their perceived intensity after unilateral brain damage. Cortex 130, 78–93. https://doi.org/10.1016/j.cortex.2020.05.015.
Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., Caclin, A., 2019. Emotional prosody in congenital amusia: impaired and spared processes. Neuropsychologia 134, 107234. https://doi.org/10.1016/j.neuropsychologia.2019.107234.
Pralus, A., Gomot, M., Graves, J., Cholvy, F., Fornoni, L., Tillmann, B., Caclin, A., 2020b. Pre-attentive processing of neutral and emotional sounds in congenital amusia. Neuroscience. https://doi.org/10.1101/2020.08.05.238204 [Preprint].
Pralus, A., Hermann, R., Cholvy, F., Aguera, P.-E., Moulin, A., Barone, P., Grimault, N., Truy, E., Tillmann, B., Caclin, A., 2021. Rapid assessment of non-verbal auditory perception in normal-hearing participants and cochlear implant users. J. Clin. Med. 10 (10), 2093. https://doi.org/10.3390/jcm10102093.
Shao, J., Wang, L., Zhang, C., 2020. Talker processing in Mandarin-speaking congenital amusics. J. Speech Lang. Hear. Res. https://doi.org/10.1044/2020_JSLHR-19-00209.
Shao, J., Zhang, C., Peng, G., Yang, Y., Wang, W.S.-Y., 2016. Effect of noise on lexical tone perception in Cantonese-speaking amusics. Interspeech 272–276. https://doi.org/10.21437/Interspeech.2016-891.
Stewart, L., 2011. EPS prize lecture: characterizing congenital amusia. Q. J. Exp. Psychol. 64 (4), 625–638. https://doi.org/10.1080/17470218.2011.552730.
Tang, C., Hamilton, L.S., Chang, E.F., 2017. Intonational speech prosody encoding in the human auditory cortex. Science 357 (6353), 797–801. https://doi.org/10.1126/science.aam8577.
Tang, W., Wang, X., Li, J., Liu, C., Dong, Q., Nan, Y., 2018. Vowel and tone recognition in quiet and in noise among Mandarin-speaking amusics. Hear. Res. 363, 62–69. https://doi.org/10.1016/j.heares.2018.03.004.
Thompson, W.F., Marin, M.M., Stewart, L., 2012. Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. Proc. Natl. Acad. Sci. USA 109 (46), 19027–19032. https://doi.org/10.1073/pnas.1210344109.
Tillmann, B., Albouy, P., Caclin, A., 2015. Congenital amusias. In: Handbook of Clinical Neurology, vol. 129. Elsevier, pp. 589–605. https://doi.org/10.1016/B978-0-444-62630-1.00033-0.
Tillmann, B., Albouy, P., Caclin, A., Bigand, E., 2014. Musical familiarity in congenital amusia: evidence from a gating paradigm. Cortex 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., 2011a. Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages. Front. Psychol. 2 https://doi.org/10.3389/fpsyg.2011.00120.
Tillmann, B., Graves, J.E., Talamini, F., Lévêque, Y., Fornoni, L., Hoarau, C., Pralus, A., Ginzburg, J., Albouy, P., Caclin, A., 2023. Auditory cortex and beyond: deficits in congenital amusia. Hear. Res. 437, 108855 https://doi.org/10.1016/j.heares.2023.108855.

Tillmann, B., Jolicœur, P., Ishihara, M., Gosselin, N., Bertrand, O., Rossetti, Y., Peretz, I., 2010. The amusic brain: lost in music, but not in space. PLoS One 5 (4), e10173. https://doi.org/10.1371/journal.pone.0010173.

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

Tillmann, B., Rusconi, E., Traube, C., Butterworth, B., Umiltà, C., Peretz, I., 2011b. Fine-grained pitch processing of music and speech in congenital amusia. J. Acoust. Soc. Am. 130 (6), 4089–4096. https://doi.org/10.1121/1.3658447.

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

Vialatte, A., Chabanat, E., Witko, A., Pisella, L., 2023. Toward the characterization of a visual form of developmental dyslexia: reduced visuo-attentional capacity for processing multiple stimuli made of separable features. Cogn. Neuropsychol. 40 (3–4), 186–213. https://doi.org/10.1080/02643294.2023.2266179.

Whiteford, K.L., Oxenham, A.J., 2018. Learning for pitch and melody discrimination in congenital amusia. Cortex 103, 164–178. https://doi.org/10.1016/j.cortex.2018.03.012.

Williamson, V.J., Cocchini, G., Stewart, L., 2011. The relationship between pitch and space in congenital amusia. Brain Cognit. 76 (1), 70–76. https://doi.org/10.1016/j.bandc.2011.02.016.

Williamson, V.J., McDonald, C., Deutsch, D., Griffiths, T.D., Stewart, L., 2010. Faster decline of pitch memory over time in congenital amusia. Adv. Cognit. Psychol. 6 (1), 15–22. https://doi.org/10.2478/v10053-008-0073-5.

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