



Syntactic processing in music and language: Parallel abnormalities observed in congenital amusia

Yanan Sun^{a,b,*}, Xuejing Lu^{b,c,d,e}, Hao Tam Ho^{f,g}, Blake W. Johnson^{a,b}, Daniela Sammler^h, William Forde Thompson^{b,c}

^a Department of Cognitive Science, Macquarie University, New South Wales 2109, Australia

^b ARC Centre of Excellence in Cognition and its Disorders, New South Wales 2109, Australia

^c Department of Psychology, Macquarie University, New South Wales 2109, Australia

^d CAS Key Laboratory of Mental Health, Institute of Psychology, Beijing 100101, China

^e Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China

^f Department of Translational Research on New Technologies in Medicine and Surgery, University of Pisa, Pisa 56126, Italy

^g School of Psychology, University of Sydney, New South Wales 2006, Australia

^h Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig 04103, Germany

ARTICLE INFO

Keywords:

Music
Language
Syntax
Congenital amusia
ERP

ABSTRACT

Evidence is accumulating that similar cognitive resources are engaged to process syntactic structure in music and language. Congenital amusia – a neurodevelopmental disorder that primarily affects music perception, including musical syntax – provides a special opportunity to understand the nature of this overlap. Using electroencephalography (EEG), we investigated whether individuals with congenital amusia have parallel deficits in processing language syntax in comparison to control participants. Twelve amusic participants (eight females) and 12 control participants (eight females) were presented melodies in one session, and spoken sentences in another session, both of which had syntactic-congruent and -incongruent stimuli. They were asked to complete a music-related and a language-related task that were irrelevant to the syntactic incongruities. Our results show that amusic participants exhibit impairments in the early stages of both music- and language-syntactic processing. Specifically, we found that two event-related potential (ERP) components – namely Early Right Anterior Negativity (ERAN) and Left Anterior Negativity (LAN), associated with music- and language-syntactic processing respectively, were absent in the amusia group. However, at later processing stages, amusics showed similar brain responses as controls to syntactic incongruities in both music and language. This was reflected in a normal N5 in response to melodies and a normal P600 to spoken sentences. Notably, amusics' parallel music- and language-syntactic impairments were not accompanied by deficits in semantic processing (indexed by normal N400 in response to semantic incongruities). Together, our findings provide further evidence for shared music and language syntactic processing, particularly at early stages of processing.

1. Introduction

In both music and language, discrete elements are combined to form larger structural units according to conventions that can be codified into a set of rules (e.g., rules of tonal structure in music and rules of morphology in language). “Syntax” has been defined broadly as a set of rules governing the combination of discrete structural elements into larger units (Asano and Boeckx, 2015). This broad definition raises the possibility that music and language draw upon shared cognitive resources for syntactic processing (Patel, 2003; see also Koelsch, 2012).

Event-related potential (ERP) studies have shown comparable

electrical brain responses during processing music-syntactic and language-syntactic violations in both early and later stages of syntactic processing. At early stages of processing (within a few hundred milliseconds latency), morpho-syntactic mismatches in sentences (e.g., gender disagreement) typically elicit a negative-going deflection with a left-hemispheric preponderance, termed the *Left Anterior Negativity* (LAN; for a review, see Friederici, 2002), which is considered an electrophysiological marker of morpho-syntactic agreement processing (Molinaro et al., 2015). Music-syntactic violations (e.g., out-of-key tones in single melodies and chord violations in harmonised melodies) elicit an early negative-going deflection with a right-hemispheric

* Corresponding author at: Department of Cognitive Science, 16 University Avenue, Macquarie University, New South Wales 2109, Australia.
E-mail address: yanan.sun@mq.edu.au (Y. Sun).

preponderance, termed the *Early Right Anterior Negativity* (ERAN), which is thought to reflect regularity-based music-syntactic processing (for a review, see Koelsch and Friederici, 2003). Koelsch et al. (2005) observed an interaction between the LAN and ERAN components when music- and language-syntactic violations occurred simultaneously. Notably, no such interactions were observed when language manipulations involved semantic incongruities, or when music manipulations involved an unexpected timbre (Koelsch et al., 2005). At later stages of processing, a positive-going deflection – the P600 – is typically elicited by morpho-syntactic violations in language (for a review, see Friederici, 2002). Converging evidence suggests that, the P600 reflects the integration, reanalysis and repair of syntactic information (Friederici, 2002). Patel et al. (1998) has shown that P600 is also elicited by violations of musical key structure and argued that this response is indistinguishable from the one elicited by violations of linguistic syntactic structure in the same participants. Finally, when the music-syntactic violations are task-irrelevant, a negativity called the N5 can be observed which supposedly reflects structure integration and meaning extraction in music (Koelsch, 2011; Koelsch et al., 2000).

Adding to the electrophysiological evidence, behavioural studies also reveal interference between music- and language-syntactic processing. For example, Slevc et al. (2009) found that reading speed for garden-path sentences was slower when combined with structurally unexpected chords than with expected chords (see also Fedorenko et al., 2009; Kunert et al., 2016; Van de Cavey and Hartsuiker, 2016).

In terms of the neural substrates underlying music- and language-syntactic processing, neuroimaging studies have shown overlapping brain regions, such as the bilateral inferior frontal gyrus (e.g., Broca's area; Janata et al., 2002; Koelsch et al., 2002c; Kunert et al., 2015; Maess et al., 2001; Tillmann et al., 2006) and superior temporal gyrus (Koelsch et al., 2002c; Sammler et al., 2013). However, it has been noted that processes associated with the same brain region are not necessarily shared, given the density of neurons within any given area (Peretz et al., 2015).

Disorders in music and language provide another avenue to examine the resource-sharing hypothesis. Music-syntactic deficits have been observed in patients with lesions in “typical language brain areas” (e.g., Patel et al., 2008; Sammler et al., 2011; but such disorders can also arise following damage to other regions, see Peretz, 1993 and Slevc et al., 2016), and in children with developmental language disorders (e.g., Jentschke et al., 2008). Language impairments have also been reported for some individuals with acquired amusia (e.g., Sarkamo et al., 2009). However, it is unclear whether individuals with developmental musical disorders exhibit deficits in both music- and language-syntactic processing.

Congenital amusia is a neurodevelopmental disorder that mainly affects music perception. Unlike typical western listeners, amusic individuals do not favour consonant over dissonant chords (Ayotte et al., 2002; Cousineau et al., 2012), and they have comparatively elevated pitch-discrimination thresholds (Ayotte et al., 2002). They also have difficulty detecting out-of-key notes in melodies in explicit tasks, suggesting reduced sensitivity to musical syntax (Peretz et al., 2002; Peretz et al., 2007). Interestingly, amusic individuals still exhibit implicit knowledge of harmonic syntax (Tillmann et al., 2012) and ERP studies suggest that they may exhibit normal brain responses to mistuned notes at early stages of processing (Mignault Goulet et al., 2012; Moreau et al., 2013; Peretz et al., 2009) but abnormal brain responses, such as an absence of early negativity, when they are asked to respond to music-syntactic mismatches (e.g., out-of-key notes; Peretz et al., 2009; Zendel et al., 2015). These explicit music-syntactic difficulties appear to be independent from their pitch discrimination deficits (Jiang et al., 2016). In other words, individuals with congenital amusia appear to have preserved brain responses to sensory violations, but abnormal brain responses to melodic syntax. Surprisingly, no investigation of congenital amusia has yet examined whether the disorder is associated with parallel deficits in music and language syntactic processing.

If there were shared mechanisms for processing syntax in music and language, then amusic individuals with music-syntactic difficulties should suffer parallel difficulties in language-syntactic processing. To test this hypothesis, we used electroencephalography (EEG) to examine brain responses to syntactic irregularities in music and language among individuals with and without congenital amusia. As a control condition, we also included language semantic irregularity as language-semantic processing is usually believed to operate independently from music-syntactic processing (Carrus et al., 2013; Kunert et al., 2016; Slevc et al., 2009).

To examine music-syntactic processing, ERPs were collected in response to syntactic violations in melodies (i.e., out-of-key notes in tone sequences). We focused on violations of melodic syntax, rather than harmonic syntax, as the latter is the most elementary instantiation of music-syntactic processing, and also because melodic syntax matches the monophonic nature of our language stimuli. A number of studies have confirmed that irregular tones in melodies elicit frontal potentials that can be interpreted as the ERAN response to music-syntactic violations (Besson and Faita, 1995; Besson and Macar, 1987; Koelsch and Jentschke, 2010; Miranda and Ullman, 2007; Paller et al., 1992). Moreover, when melodic and harmonic syntactic violations are compared directly, both elicit ERAN responses, but harmonic violations elicit additional responses that are not observed with melodic stimuli, reflecting emergent qualities that arise when individual melodic voices are combined to form a harmonic sequence (Koelsch and Jentschke, 2010). This comparison illustrates that brain responses to harmonic sequences cannot be entirely predicted from responses to the melodies that make up those harmonic sequences, corroborating earlier perceptual findings involving the same comparison (Thompson, 1993; Thompson and Cuddy, 1989, 1992).

Unlike syntactic irregularities in language, which concern violations of expectations about the function and order of words, syntactic irregularities in melody fundamentally entail unexpected acoustic information, which has the potential to complicate the interpretation of brain responses to such irregularities (Bigand et al., 2006). However, brain responses to sensory violations are evoked by an unexpected change to a sequence of elements containing a constant sensory attribute, such as pitch or loudness (Peter et al., 2010). In contrast, syntactic violations in melody do not require pitch (or other sensory attributes) to be held constant in a sequence, because syntax operates at a more abstract level that is determined by the implied tonal hierarchy.

In this investigation, we re-examined whether individuals with congenital amusia exhibit typical brain responses to violations of melodic syntax, while also investigating whether they exhibit typical brain responses to language-syntactic irregularities. We hypothesised that, in comparison to the control group, amusic individuals would exhibit abnormal brain responses to both music-syntactic violations and language-syntactic irregularities if processing music and language syntax involves shared cognitive mechanisms. However, we expected that amusic and control groups would exhibit normal brain responses to non-syntactic unexpected events in music and language.

2. Materials and methods

2.1. Assessment of congenital amusia

In the present study, amusic participants were identified using a screening method based on the three pitch-related subtests (Scale, Contour and Interval) from the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al., 2003) with an aggregate accuracy rate of 72.22% being the cutoff (i.e., 65 out of 90 points; Liu et al., 2010; Sun et al., 2017; Thompson et al., 2012). The ability to detect changes in melodic pitch, assessed by the three subtests that we employed, is fundamental to the processing melodic syntax, which was the focus of our investigation. Given that the cutoff based on the percentage of correct responses is subject to response bias, which may lead to

misclassification (Henry and McAuley, 2013; Pfeifer and Hamann, 2015), we also calculated the corresponding d' score of the aggregate accuracy rate, based on the hit and false alarm rate obtained on these three subtests for each participant (Stanislaw and Todorov, 1999).

2.2. Participants

Twelve monolingual Australian English speakers aged 18–37 years with congenital amusia (eight females) and 12 controls (eight females) participated in this study. In our sample, no overlap with regard to d' scores was found between the amusia group ($M = 1.02$, $SD = 0.46$, range [0.23–1.87]) and the control group ($M = 2.82$, $SD = 0.54$, range [2.12–3.71]), validating our group assignment.

To evaluate participants' sensitivity to music-syntax, an out-of-key detection task was administered to all participants. The task was taken from Peretz et al. (2008) and consisted of twelve melodies and twelve out-of-key versions of those melodies, drawn from the Scale subtest of MBEA. On each trial, participants were presented with a single melody and they judged whether the melody contained a “sour” or “strange” note (i.e., outside of the implied key). The amusic participants performed worse on the out-of-key detection task compared with the control participants [for amusics, mean of $d' = 1.63$, $SE = 0.17$; for controls, mean of $d' = 2.57$, $SE = 0.18$; $t(22) = -3.92$, $p < 0.001$].

All participants were recruited from a pool of participants who were all at a university level of education. They were all right-handed, as determined by the Edinburgh Handedness Inventory (Oldfield, 1971). They also had normal hearing (< 30 dB) in both ears at the frequencies of 0.25, 0.5, 1, 2, 4, and 8 kHz, which was confirmed using an Otovation Amplitude T3 series audiometer (Otovation LLC, PA, United States). Individuals with dyslexia were excluded from the investigation, and previous research confirms that amusic and control participants have similar general linguistic skills (Sun et al., 2017). No participant reported neurological or psychiatric disorders. Table 1 provides an overview of the participants' characteristics. Written informed consent was provided from all participants. The Macquarie University Ethics Committee approved the research protocol.

2.3. Stimuli

To investigate amusics' and controls' ability to process syntactically congruent and incongruent melodies, we created 80 melodies in C major with a piano timbre. Each melody consisted of five consecutive notes. All melodies were compatible with a C major scale and considered music-syntactically congruent (see Fig. 1A). In order to create

Table 1
Summary of amusic and control participants' characteristics and test scores.

	Amusics (n = 12)	Controls (n = 12)	t-tests
	Mean (SE)	Mean (SE)	t
Age (years)	21.43 (1.57)	20.96 (0.93)	0.23
Education (years)	14.38 (0.55)	14.08 (0.36)	0.45
Musical Training (years)	0.33 (0.26)	0.79 (0.31)	-1.12
Melodic MBEA (%)			
Scale	76.39 (3.16)	94.44 (1.44)	-5.19***
Contour	60.28 (1.86)	90.55 (2.04)	-10.97***
Interval	62.78 (2.04)	83.05 (2.41)	-6.42***
Global score	66.48 (1.66)	89.35 (1.52)	-10.15***
Out-of-key detection (d')	1.63 (0.17)	2.57 (0.18)	-3.92***

Global score indicates the average of the individual scores on the three subtests (Scale, Contour and Interval) of MBEA. Subtest scores and global scores are expressed in percentages. The ability to detect out-of-key notes was evaluated using d' . Group differences were assessed using independent samples t -tests (two-tailed).

*** Denotes $p < 0.001$.

music-syntactically incongruent melodies, the final tones of the original 80 melodies were shifted up or down until they were no longer compatible with a C major scale (out of key), yielding 80 syntactically incongruent melodies. The pitch change direction from the penultimate tone to the final tone remained identical to their syntactically congruent counterparts. The pitch interval between these two tones were also controlled to ensure that there was no significant difference between the congruent and incongruent melodies [for congruent melodies, $M = 3.30$ semitones, $SE = 0.23$ semitones; for incongruent melodies, $M = 3.41$ semitones, $SE = 0.23$ semitones; $t(158) = -0.35$, $p = 0.728$]. All tones fell within the pitch range of G3 to C5. Each of the first four notes was 600 ms in duration and the final note was 1200 ms.

Using the same procedure, additional 20 congruent and 20 incongruent probe melodies were created. Each of the probes included a single tone with a guitar timbre instead of a piano timbre. The deviant tone could occur at any position except for the first position in the melody. These probe melodies were included to ensure that participants paid attention to all melodies (see Procedures subsection for details). Brain responses to probe stimuli were excluded from the ERP analyses of congruent and incongruent melodies. All the 200 melodies were digitally generated using GarageBand 6.0.5 (Apple Inc., CA, United States).

To test participants' language processing abilities, 80 five-word morpho-syntactically and semantically congruent sentences were originally created. Each sentence had a fixed syntactic structure with the general form “[someone] is [doing] [one/two] [thing/things]”. The final word in all original sentences was modified to generate 80 morpho-syntactically incongruent but semantically congruent sentences, and another 80 morpho-syntactically congruent but semantically incongruent sentences (for stimulus examples, see Fig. 1B). To prevent the syntactic correctness being solely determined by the presence/absence of an “s”, each of the three sentence-types comprised 40 sentences ending with a singular noun and 40 sentences ending with a plural noun. An additional 80 morpho-syntactically and semantically congruent sentences were included as filler sentences to ensure the whole stimuli pool had equal number of congruent and incongruent sentences. Congruent sentences were spoken in a natural manner by a female monolingual Australian English speaker, and recorded for stimulus presentation. Incongruent versions of these sentences were created by splicing the final words in the congruent sentences using the computer software Praat (Boersma and Weenink, 2014). The resultant 320 spoken sentences ranged from 1.91 s to 2.88 s in duration ($M = 2.57$ s, $SD = 0.17$ s).

2.4. Procedures

The experiment consisted of a music session and a language session, both of which were completed by all participants in one day. The music and language session lasted approximately 30 and 40 min, respectively. The order of the sessions was counterbalanced across participants. All stimuli were presented via insert earphones (Model ER-30, Etymotic Research Inc., IL, United States) with the intensity level fixed at 80 dB SPL for all participants.

In the music session, participants were presented with 200 melodies in random order. To ensure that ERPs reflected music-syntactic processing but not explicit decisional processes related to our music-syntactic manipulations, participants were not informed of the music-syntactic incongruities. Instead, they were only instructed to detect the timbre-deviants (see previous section), which only occurred in the probe melodies and were excluded from the analysis. Responses were given via a button press on a response pad (HHSC-2 × 2, Current Designs Inc., PA, USA). Prior to testing, participants were presented the C major scale five times in both forward and reverse order to induce a strong C major context.

In the language session, all participants listened to 320 sentences played in random order. Similar to the music session, participants were







A Music	B Language
 <p>Music-syntactic congruent melody</p>	 Syntactically congruent and semantically congruent sentences Singular Linda is feeding one pig. Plural Linda is feeding two pigs.
 <p>Music-syntactic incongruent melody</p>	 Syntactically incongruent but semantically congruent sentences Singular Linda is feeding two pig. Plural Linda is feeding one pigs.
	 Syntactically congruent but semantically incongruent sentences Singular Linda is feeding one kite. Plural Linda is feeding two kites.
	 Filler sentences Singular Linda is flying one kite. Plural Linda is flying two kites.

Fig. 1. Stimuli used in the music and language EEG sessions. (A) Two example stimuli (congruent and incongruent) presented in the music session. (B) Variants of an example stimulus presented in the language session. The critical positions are marked with grey-shading.

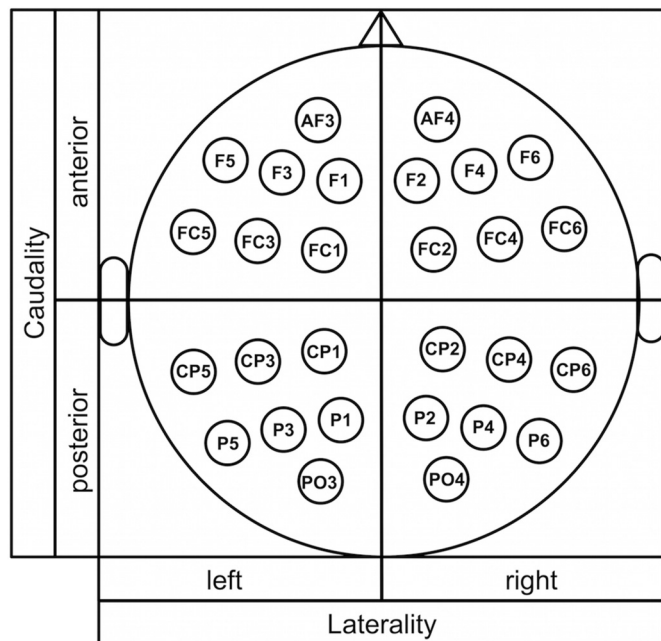


Fig. 2. Electrode ROIs used for statistical analyses.

not informed about the morpho-syntactic incongruities or semantic incongruities. Instead, they were only instructed to listen to all sentences attentively. To check that participants were paying attention, sixteen probe questions were presented randomly during the session to query the content of a preceding filler sentence. Different questions probed participants' understanding of the subject, verb or object of the sentences (e.g., "Who is flying one kite?", "What is Linda doing with one kite?", and "What is Linda flying?"). Participants did not know in advance which part of the sentence they would be asked about, and therefore needed to attend to the entire sentence in each trial. Participants' verbal responses were recorded and subsequently coded for statistical analysis.

2.5. EEG data acquisition and processing

The EEG was recorded at 1 kHz using a BrainAmp MR amplifier (BrainProducts GmbH, Gilching, Germany). Participants wore an EEG cap with 63 Ag/AgCl electrodes, one of which was used to monitor eye

movements and blinks. Electrode impedances were kept below 5 k Ω and a band-pass filter of 0.03–200 Hz was applied online. The FCz electrode was used as an online reference.

The recorded EEG data were analysed offline using EEGLAB 13.5.4b (Delorme and Makeig, 2004) in MATLAB 8.5 (MathWorks Inc., MA, United States). All data were resampled to 500 Hz and filtered with a 0.1 Hz high-pass windowed sinc FIR filter with a Blackman window and a transition bandwidth of 0.15 Hz (the corresponding filter length was 18,334; Widmann and Schroger, 2012). Using the TrimOutlier plugin in the EEGLAB toolbox, we identified and removed noisy channels that had standard deviations equal to or greater than 100 μ V. Afterwards, all data was re-referenced to the average reference. The data were first segmented into long epochs with trial lengths of –500 to 3900 ms relative to the first tone onset in melodies and –500 to 3500 ms relative to the first word onset in sentences. Epochs that contained probe melodies, filler sentences, button presses, and gross artifacts (< 5% of trials) were excluded from subsequent data processing. To identify and remove eye movement and blink artifacts, we conducted an independent components analysis using the *runica* algorithm implemented in the EEGLAB toolbox. Subsequently, bad channels were interpolated using spherical interpolation. The EEG data were then time-locked to the onset of the final note of the melodies for the music-syntactic trials; the suffix "s" that determined whether the nouns were singular or plural for the language-syntactic trials; and the final word in the sentences for the language semantic trials. Finally, shorter epochs ranging from –200 to 1000 ms were extracted and baseline corrected, using a 200 ms pre-stimulus interval.

2.6. Statistical analyses

As shown in Fig. 2, the scalp electrodes were grouped into four regions of interest (ROIs) a priori: left-anterior, right-anterior, left-posterior and right-posterior. Visual inspection of the grand averages revealed five large deflections in the ERPs at around (i) 130–250 ms and (ii) 500–650 ms in the music condition; around (iii) 120–250 ms and (iv) 500–650 ms in the language syntactic condition; and (v) 300–500 ms in the language semantic condition. The time windows of these deflections were in line with the time windows of the (i) Early Right Anterior Negativity (ERAN; e.g., Koelsch et al., 2000), (ii) N5 (e.g., Koelsch et al., 2000), (iii) Left Anterior Negativity (LAN; e.g., Hasting et al., 2007), (iv) P600 (Gouvea et al., 2010), and (v) N400 (Kutas and Federmeier, 2000).

An anterior negativity and a posterior positivity can be seen in the

400–1000 ms time windows for music- and language-syntactic conditions (see the details in the [Results](#) section). In previous studies, a P600 response is usually elicited when participants explicitly respond to a syntactic violation in music (Patel et al., 1998). Conversely, N5 rather than P600 is observed when attention is not drawn to these musical events (Koelsch et al., 2000). In our music experiment, participants were not instructed to pay attention to the syntactic structure. Thus, the late component elicited by the music-syntactic incongruities is most likely to an N5. Similarly, based on previous research, the late component elicited by the language-syntactic incongruities in our study is most likely a P600.

For each ERP component (i–v), mean amplitudes of the corresponding time window were computed and then entered into repeated measures ANOVAs with one between-subject factor Group (amusia vs. control) and three within-subject factors: Congruency (congruent vs. incongruent), Laterality (left vs. right hemisphere) and Caudality (anterior vs. posterior). Whenever an interaction involving Group and Congruency was significant, follow-up ANOVAs were conducted for further analysis. Effect size was estimated using generalised eta-squared (η^2). We will report in detail only significant main effects and interactions of interest for each time window. For a full list of the statistical results, see [Table 2](#).

To examine participants' performance on the timbre-deviant detection task in the music session, we computed d' scores [$d' = z(\text{hit rate}) - z(\text{false alarm rate})$; Stanislaw and Todorov, 1999] for each individual and conducted one-sample t -tests for each group.

3. Results

3.1. Response to probe trials

Participants' performance in the probe trials confirmed that they attended to melodic and language stimuli. For the melody session, the results of the timbre-deviant detection task indicate that amusic and control participants performed above chance level ($d' = 0$) [controls: $M = 3.21$, $SE = 0.26$, $t(11) = 11.23$, $p < 0.001$; amusics: $M = 2.71$, $SE = 0.29$, $t(11) = 10.27$, $p < 0.001$]. An independent-sample t -test showed that the two groups did not significantly differ in their ability to detect timbre changes [$t(22) = 1.28$, $p = 0.21$]. In the language task, twenty out of twenty-four participants answered all sixteen probe questions correctly. Two amusic and two control participants made one error each. We informally interviewed a sample of the participants to determine whether they noticed syntactic violations even though their attention was directed towards a non-syntactic change. Participants generally reported noticing an anomaly in the stimuli, suggesting that syntactic processing occurred automatically and that violations in syntax captured attention to some degree.

3.2. ERP results

3.2.1. ERAN (130–250 ms)

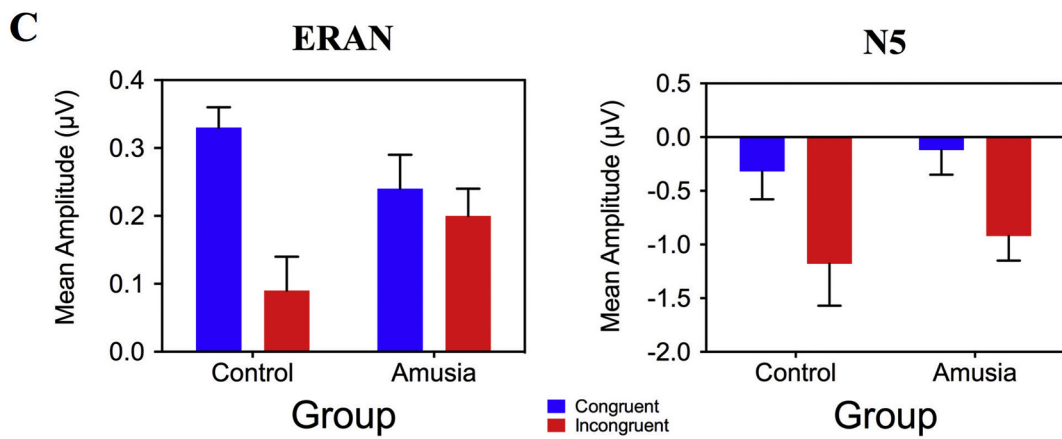
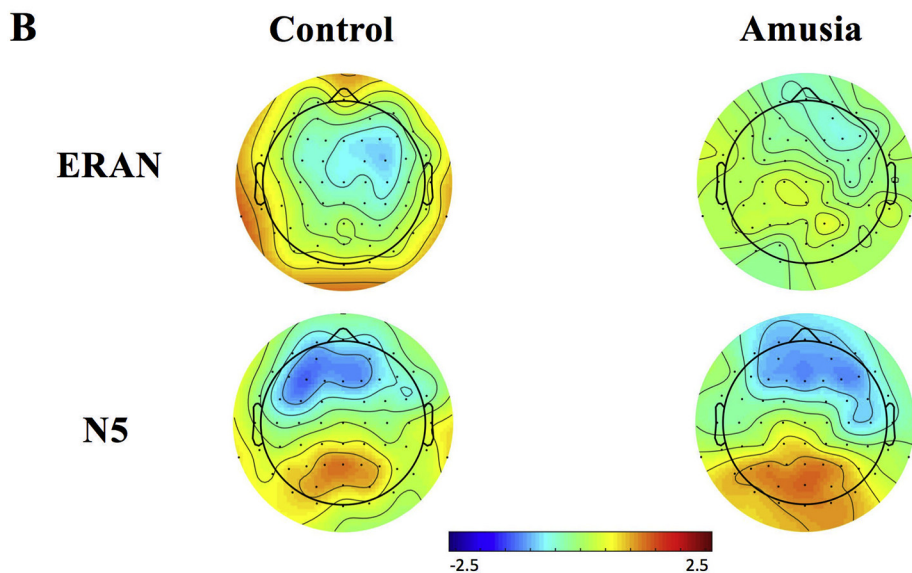
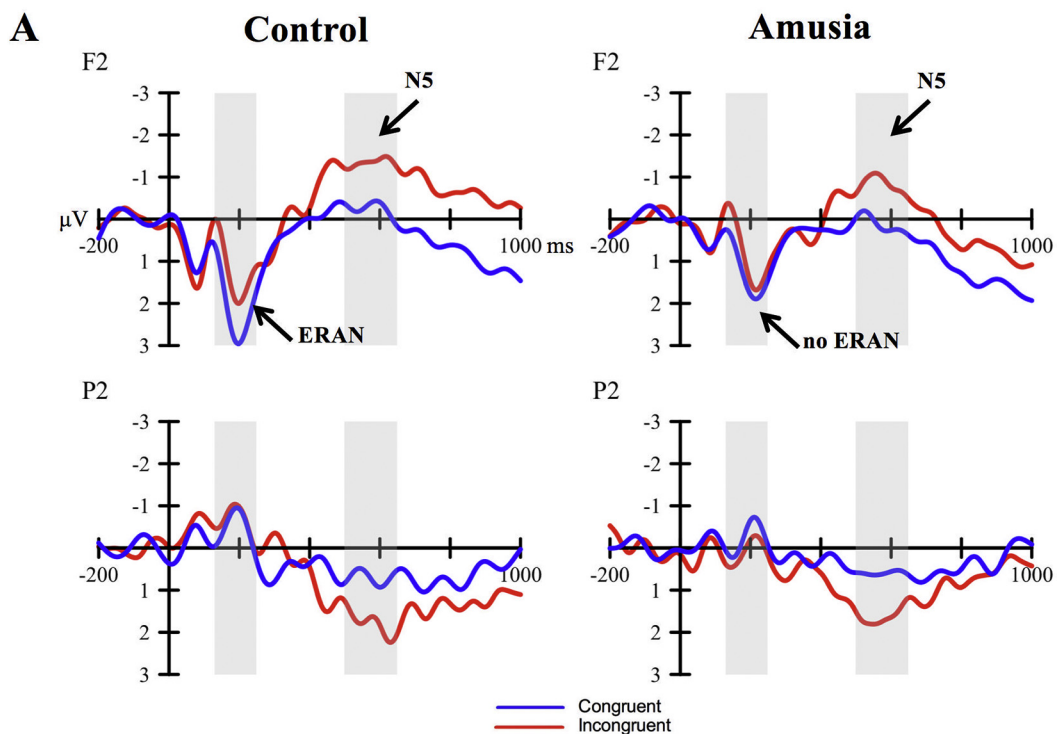
Music-syntactic incongruities evoked a negative-going ERP response in the control but not the amusia group. This negativity is likely the early right anterior negativity (ERAN; e.g., Koelsch et al., 2000; see [Fig. 3](#)). Confirming our observation, the ANOVA yielded a significant Group by Congruency interaction [$F(1, 22) = 7.28$, $p = 0.013$, $\eta^2 = 0.004$]. Follow-up ANOVAs conducted separately for the amusia and control groups pointed to a clear difference between the incongruent and congruent condition in the control group [incongruent: $M = 0.09 \mu\text{V}$, $SE = 0.05 \mu\text{V}$; congruent: $M = 0.33 \mu\text{V}$, $SE = 0.03 \mu\text{V}$; $F(1, 11) = 14.20$, $p = 0.003$, $\eta^2 = 0.394$]. This contrast was not significant in the amusia group [incongruent: $M = 0.20 \mu\text{V}$, $SE = 0.04 \mu\text{V}$; congruent: $M = 0.24 \mu\text{V}$, $SE = 0.05 \mu\text{V}$; $F(1, 11) = 1.27$, $p = 0.283$, $\eta^2 = 0.015$; see also [Fig. 3C](#), left panel]. We also analysed the effect of Congruency for amusic and control participants within the right anterior ROI only where ERAN typically distributes. The results confirmed

Table 2
ANOVA results for ERP components.

Effects	Music-syntax			Language-syntax			Language-semantics								
	ERAN (130–250 ms)			LAN (120–250 ms)			P600 (500–650 ms)			N400 (300–500 ms) ^a					
	F (1, 22)	p	η^2	F (1, 22)	p	η^2	F (1, 22)	p	η^2	F (1, 22)	p	η^2			
G	0.02	0.889	< 0.001	1.10	0.304	0.001	0.02	0.900	< 0.001	1.16	0.293	0.003	0.30	0.589	0.001
L	1.71	0.204	0.006	10.23	0.004	0.036	2.83	0.107	0.018	0.20	0.659	0.001	1.84	0.190	0.013
C	38.38	< 0.001	0.551	16.12	< 0.001	0.322	7.92	0.010	0.178	1.06	0.313	0.029	3.65	0.069	0.069
Co	14.64	< 0.001	0.009	2.03	0.168	0.001	7.46	0.012	0.008	0.28	0.600	< 0.001	41.42	< 0.001	0.039
G × L	0.28	0.602	0.001	0.02	0.901	< 0.001	0.09	0.761	0.001	0.88	0.359	0.003	0.62	0.440	0.004
G × C	3.04	0.095	0.076	0.27	0.607	0.008	0.171	0.171	0.052	1.10	0.305	0.030	0.18	0.679	0.004
G × Co	7.28	0.013	0.004	0.23	0.640	< 0.001	0.76	0.394	0.001	1.24	0.278	0.001	0.33	0.572	< 0.001
L × C	0.14	0.708	0.000	3.45	0.077	0.004	1.83	0.190	0.002	2.56	0.124	0.003	5.41	0.030	0.008
L × Co	2.63	0.119	0.008	0.19	0.668	< 0.001	1.29	0.269	0.001	1.23	0.279	0.002	1.10	0.307	0.002
C × Co	3.17	0.089	0.023	17.68	< 0.001	0.129	3.47	0.076	0.022	20.68	< 0.001	0.121	0.21	0.654	0.002
G × L × C	0.43	0.602	0.001	0.98	0.353	< 0.001	1.10	0.305	0.002	1.45	0.241	0.002	0.11	0.739	< 0.001
G × L × Co	0.24	0.627	0.001	1.98	0.173	0.002	4.68	0.042	0.005	0.06	0.802	< 0.001	0.06	0.816	< 0.001
G × C × Co	0.03	0.861	< 0.001	0.01	0.984	< 0.001	0.05	0.988	< 0.001	0.52	0.478	0.003	0.01	0.983	< 0.001
L × C × Co	2.00	0.060	0.002	1.51	0.232	0.001	0.04	0.847	< 0.001	0.03	0.865	< 0.001	0.34	0.568	< 0.001
G × L × C × Co	0.28	0.602	0.000	1.36	0.257	0.001	4.66	0.042	0.002	0.03	0.859	< 0.001	0.13	0.723	< 0.001

Bold values indicate significant results ($p < 0.05$). G = Group, L = Laterality, C = Caudality, and Co = Congruency.

^a We also evaluated the N400 with a more classical ROI reported in the literature including the electrodes in central area (FC3, FC4, C3, C4, CP3, CP4, FC1, FC2, C1, Cz, C2, CP1, CP2, P1, Pz, P2). Similar results were found, where the congruency effect was significant [$F(1, 22) = 39.23$, $p < 0.001$, $\eta^2 = 0.267$], but neither significant group effect nor interaction between congruency and group was observed.



(caption on next page)

Fig. 3. Music-syntactic results. (A) Grand-average ERPs at electrode F2 (top) and P2 (bottom) in 12 controls (left) and 12 amusics (right), are time-locked to the onset of music-syntactically congruent (blue line) or incongruent tones (red line). ERAN and N5 are indicated by arrows, and their time-windows used for statistical analyses are marked by grey-shaded boxes. These lines are smoothed using spline interpolation for display purpose. (B) The scalp topographies of ERAN (top) and N5 (bottom) represent the amplitude difference between the music-syntactically incongruent and congruent conditions in the time windows used for statistical analyses. (C) The bar charts show the mean amplitudes in response to music-syntactically congruent (blue bar) and incongruent tones (red bar), over four ROIs (left-anterior, left-posterior, right-anterior, and right-posterior) for ERAN, and over the anterior ROIs (left-anterior and right-anterior) for N5. Each error bar represents 1 SE.

that there was a significant effect of Congruency for control participants ($p = 0.018$) but not for amusic participants ($p = 0.184$), suggesting that the two groups responded differentially to congruency.

3.2.2. N5 (500–650 ms)

In the later time window (500–650 ms), both amusia and control groups exhibited a clear ERP difference between syntactic-congruent and -incongruent tones (see Fig. 3A), with negative polarity at anterior and positive polarity at posterior electrode sites. Latency and topography of this response are consistent with that of the N5 component described by Koelsch (2005). In line with our observations, the ANOVA confirmed that there was no effect of Group (all p 's > 0.05, see also Table 2). Only the interaction between Caudality and Congruency was significant [$F(1, 22) = 17.68, p < 0.001, \eta^2 = 0.129$]. Separate ANOVAs conducted for anterior and posterior electrodes revealed significant Congruency effects at both anterior and posterior electrode sites [for anterior, $F(1, 23) = 22.23, p < 0.001, \eta^2 = 0.158$; for posterior, $F(1, 23) = 12.86, p = 0.002, \eta^2 = 0.132$] but with different patterns. In the anterior region, music-syntactically incongruent tones elicited a larger negativity ($M = -1.05 \mu\text{V}, SE = 0.22 \mu\text{V}$) than the congruent tones ($M = -0.22 \mu\text{V}, SE = 0.17 \mu\text{V}$), while the opposite was true in the posterior region, (incongruent tones: $M = 1.06 \mu\text{V}, SE = 0.21 \mu\text{V}$; congruent tones: $M = 0.38 \mu\text{V}, SE = 0.15 \mu\text{V}$; see Fig. 3C, right panel).

3.2.3. LAN (120–250 ms)

The control group exhibited a large negativity in response to the morpho-syntactic incongruence (see Fig. 4A, left panel) that was most pronounced over anterior electrode sites (see Fig. 4B, left panel), which is typical of a LAN (Steinhauer and Drury, 2012). No such component was elicited for the amusia group. Correspondingly, the ANOVA results yielded a significant four-way interaction between Group, Congruency, Laterality and Caudality [$F(1, 22) = 4.66, p = 0.042, \eta^2 = 0.002$]. Separate ANOVAs conducted for the left and right hemispheres revealed significant main effects of Caudality [$F(1, 22) = 8.40, p = 0.008, \eta^2 = 0.210$] and Congruency [$F(1, 22) = 8.68, p = 0.007, \eta^2 = 0.016$] as well as a significant interaction between Group and Congruency [$F(1, 22) = 5.19, p = 0.033, \eta^2 = 0.01$] for the left hemisphere. A closer examination of the Group by Congruency interaction revealed that the Congruency effect was significant in the control group [$F(1, 11) = 15.03, p = 0.003, \eta^2 = 0.341$], but not in the amusia group [$F(1, 11) = 0.20, p = 0.660, \eta^2 < 0.001$]. For the control group, the morpho-syntactically incongruent condition elicited a more negative-going deflection ($M = 0.07 \mu\text{V}, SE = 0.06 \mu\text{V}$) than the morpho-syntactically congruent condition ($M = 0.38 \mu\text{V}, SE = 0.07 \mu\text{V}$) in the left hemisphere (refer to Fig. 4C, left panel). In contrast, this difference was not observed in the amusia group (for the incongruent condition, $M = 0.23 \mu\text{V}, SE = 0.08 \mu\text{V}$; and for the congruent condition, $M = 0.27 \mu\text{V}, SE = 0.06 \mu\text{V}$). For the right hemisphere, the ANOVA yielded no significant effect of Group or Congruency (all p 's > 0.05, see also Table 2). Again, we analysed the effect of Congruency for amusic and control participants within the left anterior ROI only where LAN typically distributes. The results confirmed that there was a significant effect of Congruency for control participants ($p = 0.026$) but not for amusic participants ($p = 0.110$), suggesting that the two groups responded differentially to congruency.

3.2.4. P600 (500–650 ms)

In the later time window, both amusia and control groups showed a larger positivity for morpho-syntactically incongruent (see Fig. 4A and B; $M = 0.61 \mu\text{V}, SE = 0.11 \mu\text{V}$) than for syntactically congruent sentences ($M = -0.07 \mu\text{V}, SE = 0.12 \mu\text{V}$) at posterior electrode sites. At the same time, morpho-syntactically incongruent words ($M = 0.26 \mu\text{V}, SE = 0.11 \mu\text{V}$) also elicited a more negative-going deflection than syntactically congruent words at anterior electrodes ($M = 0.88 \mu\text{V}, SE = 0.14 \mu\text{V}$). These characteristics are consistent with that of the P600 (Gouvea et al., 2010). The statistical results confirmed that there was no significant Group effects or interactions involving Group and Congruency (all p 's > 0.05). Furthermore, the results showed that the interaction between Caudality and Congruency was significant [$F(1, 22) = 20.68, p < 0.001, \eta^2 = 0.121$]. Follow-up tests were conducted separately for the anterior and posterior areas, revealing a significant effect of Congruency in both regions [anterior: $F(1, 23) = 15.26, p < 0.001, \eta^2 = 0.118$; posterior: $F(1, 23) = 23.44, p < 0.001, \eta^2 = 0.156$; refer to Fig. 4C, right panel], but in opposite directions.

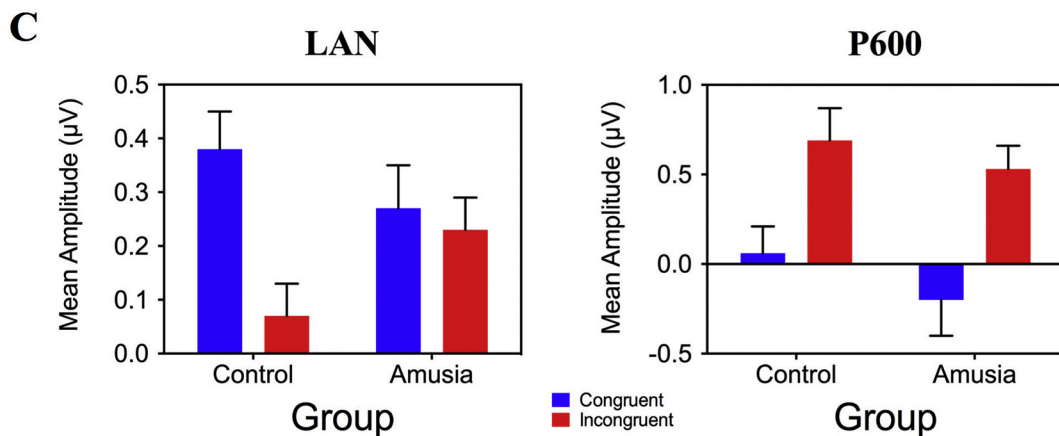
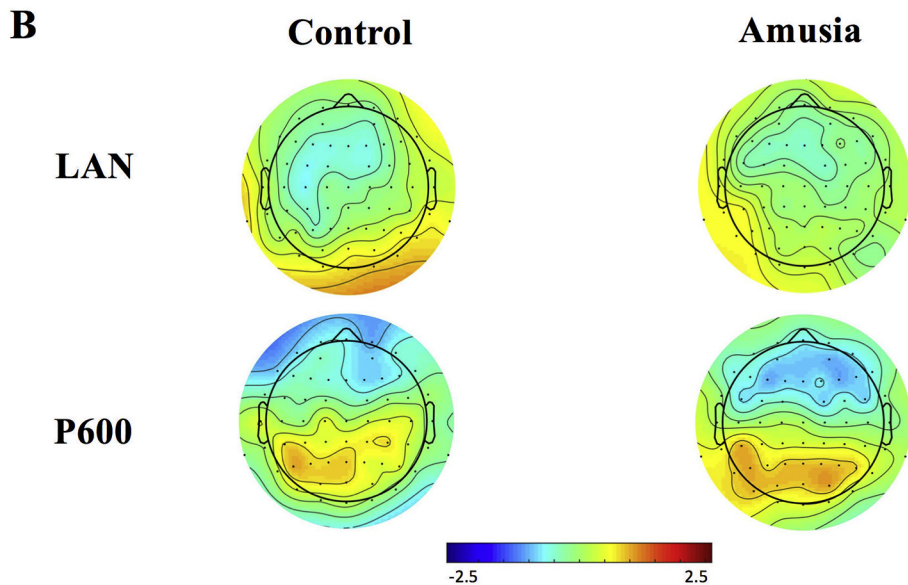
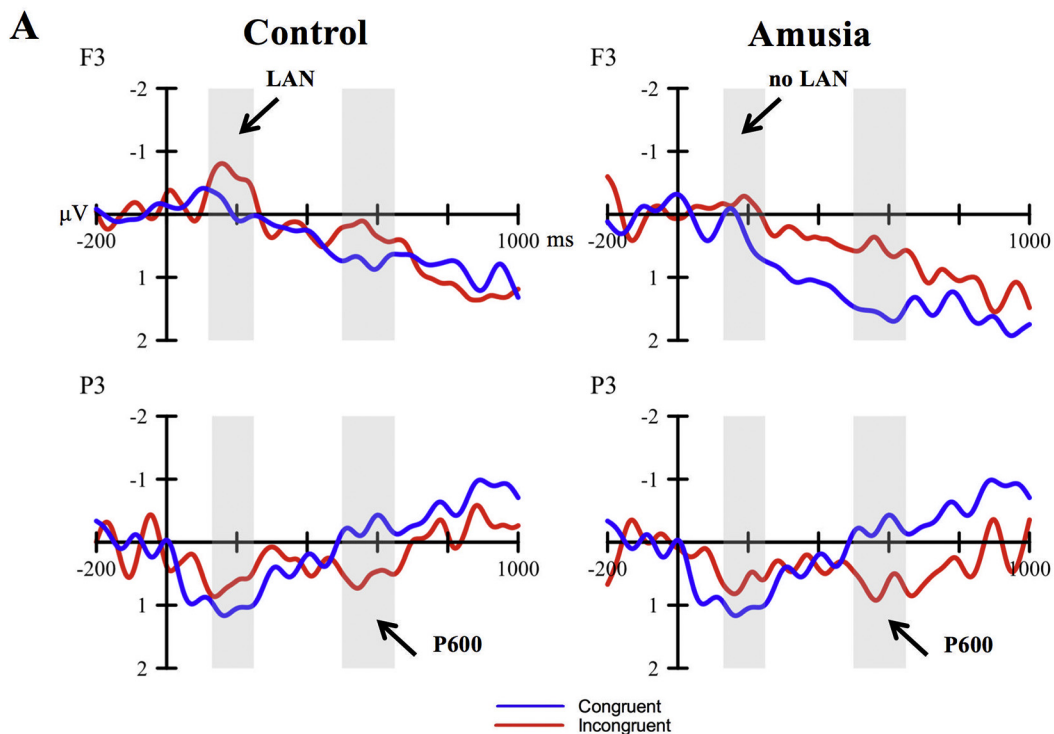
3.2.5. N400 (300–500 ms)

In order to test whether early deficits in language processing exhibited by amusics are restricted to syntax or generalize to semantics, we compared the ERPs to semantically congruent and incongruent words. We found that both amusia and control groups showed similar negative-going deflections to semantic incongruities, between 300 and 500 ms post stimulus onset. As shown in Fig. 5, this negativity has a broadly distributed topography, which is typical of the N400 (see Fig. 5A and B). In line with our observation, the ANOVA yielded no significant main effect of Group ($p > 0.05$). Instead, there was a significant main effect of Congruency [$F(1, 22) = 41.42, p < 0.001, \eta^2 = 0.039$]. Mean amplitudes revealed that this effect was due to a larger negativity elicited by semantically incongruent words ($M = -0.41 \mu\text{V}, SE = 0.05 \mu\text{V}$) as compared with semantically congruent words ($M = -0.14 \mu\text{V}, SE = 0.05 \mu\text{V}$; refer to Fig. 5C).

4. Discussion

We examined the hypothesis that individuals with congenital amusia have parallel deficits for processing syntactic structure in music and language. In separate sessions, amusic and control participants were presented with sentences and melodies containing syntactically congruent or incongruent words or tones. We reasoned that if the same cognitive resources were recruited for music- and language morpho-syntactic processing, then amusics should show abnormal brain responses, not only to music-syntactic incongruities, but also to morpho-syntactic incongruities. Consistent with these predictions, amusic participants displayed reduced ERP responses to both music-syntactic and morpho-syntactic incongruities in the early processing stage, while their brain responses were similar to those of the control participants in the late processing stage. Furthermore, amusics exhibited normal processing of semantic irregularities, as reflected by a typical N400, suggesting an impairment that is specific to early syntactic processing.

The ERAN – an ERP associated with early music-syntactic processing – was absent in the amusia group. This result is consistent with previous evidence that amusics exhibit a reduced or no early anterior negativity in response to unexpected notes (Braun et al., 2008; Omigie et al., 2013; but see Zendel et al., 2015). Specifically, Braun et al. (2008) investigated tune deafness (which is comparable to congenital amusia but



(caption on next page)

Fig. 4. Language-syntactic results. (A) Grand-average ERPs at electrode F3 (top) and P3 (bottom) in 12 controls (left) and 12 amusics (right), are time-locked to the onset of language-syntactically congruent (blue line) or incongruent condition (red line). LAN and P600 are indicated by arrows, and their time-windows used for statistical analyses are marked by grey-shaded boxes. These lines are smoothed using spline interpolation for display purposes. (B) The scalp topographies of LAN (top) and P600 (bottom) represent the amplitude difference between the language-syntactically incongruent and congruent conditions in the time windows used for statistical analyses. (C) The bar charts show the mean amplitudes in response to language-syntactically congruent (blue bar) and incongruent condition (red bar), over the left ROIs (left-anterior and left-posterior) for LAN, and over the posterior ROIs (left-posterior and right-posterior) for P600. Each error bar represents 1 SE.

is usually diagnosed using a different test) and found that incorrect tones occurring at the end of popular melodies elicited a mismatch negativity (MMN) in the control group but in the tune deaf group. Similarly, Omigie et al. (2013) reported an reduction of the MMN in amusics' response to less expected notes in melodies. There are some of inconsistencies in the terminology of MMN and ERAN, but the MMN reported in these studies, like the ERAN, reflects musical syntactic processing (Näätänen et al., 2007). These combined findings suggest that music-syntactic deficits in amusics start at early processing stages. Although these findings do not align with the pattern of results described in a similar study by Zendel et al. (2015), the ERAN in their

study was attenuated in the amusic group relative to the control group, and may have been an artefact of the P1 and N1 (refer to Fig. 4C and D in Zendel et al., 2015). Further research is needed to resolve this apparent discrepancy.

At later stages of processing, we observed no evidence of impairment: both amusic and control participants exhibited a similar N5 response. It should be emphasized, however, that we measured *implicit* processing of syntax by asking participants to focus on timbre deviants. The N5 is thought to reflect implicit music-syntactic integration and has also been associated with the processing of musical meaning (Koelsch, 2005; Koelsch et al., 2002a; Steinbeis and Koelsch, 2008). Thus,

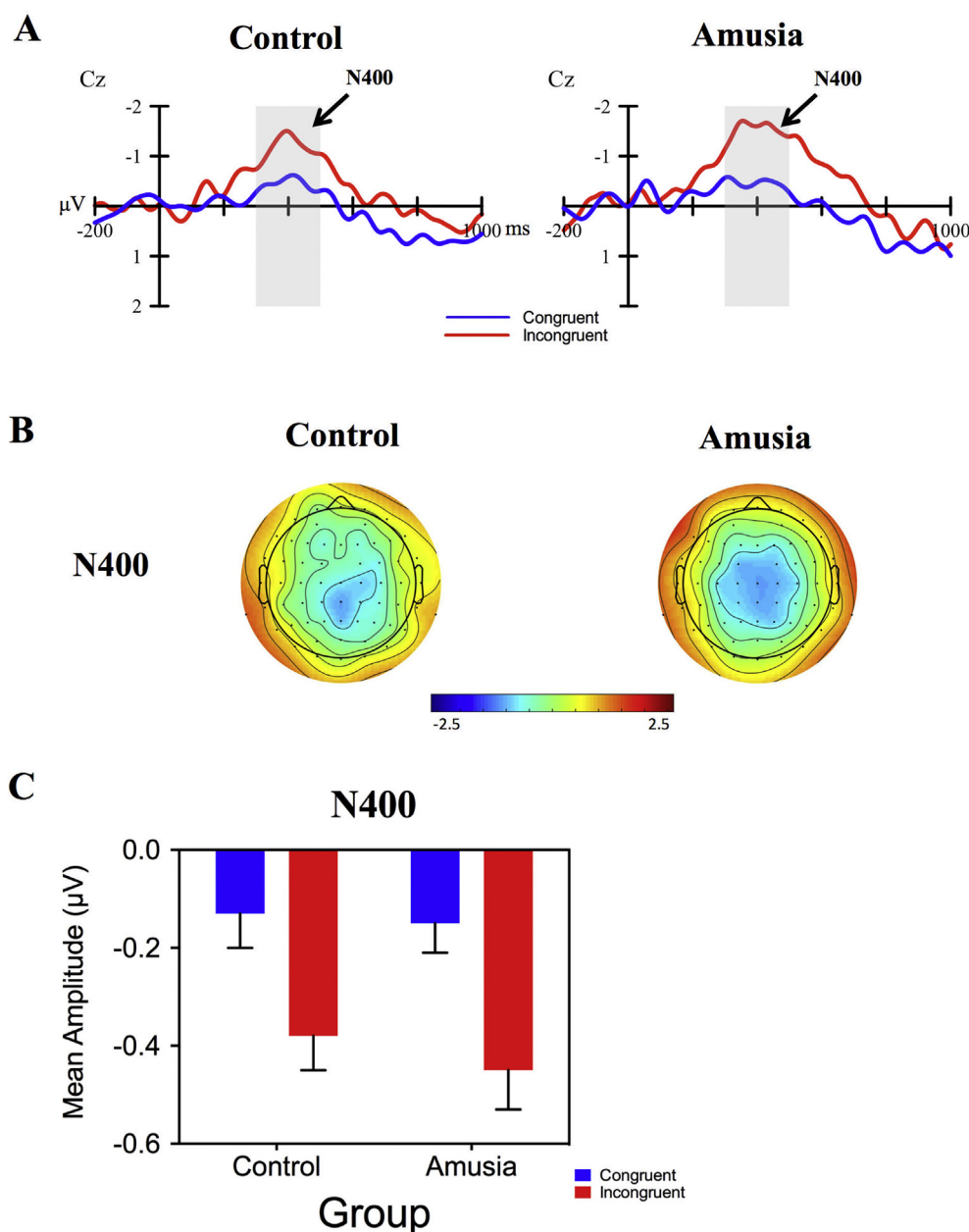


Fig. 5. Language semantic results. (A) Grand-average ERPs at electrode Cz in 12 controls (left) and 12 amusics (right), are time-locked to the onset of language-semantically congruent (blue line) or incongruent words (red line). N400 is indicated by an arrow, and its time-windows used for statistical analyses are marked by a grey-shaded box. These lines are smoothed using spline interpolation for display purpose. (B) The scalp topography of N400 represents the amplitude difference between the language-semantically incongruent and congruent conditions in the time windows used for statistical analyses. (C) The bar charts show the mean amplitudes in response to language-semantically congruent (blue bar) and incongruent words (red bar), over four ROIs (left-anterior, left-posterior, right-anterior, and right-posterior). Each error bar represents 1 SE.

explicit, conscious processing of syntax may still be disrupted at later stages. Indeed, when amusic participants are asked to make explicit judgments of syntactic congruity, the P600 (which typically reflects conscious syntactic processing) can be abnormal (Peretz et al., 2009; Zendel et al., 2015). In the current study, an impairment of conscious syntactic processing among amusic listeners was manifested behaviourally in their low average scores on the out-of-key detection task (see Table 1).

One other factor may help to account for previous findings of impaired syntactic processing at a late stage of processing. In the studies reported by Peretz et al. (2009) and Zendel et al. (2015), music-syntactic anomalies were followed by another tone. In our study, they were presented in sequence-final position. Hence, the observed late component (P600) in the former studies may reflect the processing of *relationships* between an out-of-key note and a subsequent (in-key) note – a process known as “anchoring” (Bharucha, 1984). In contrast, the late component reported in the present study may reflect the implicit integration of an out-of-key note into an established tonal schema. Taken together, the evidence suggests that amusics may have an impairment in melodic anchoring at early stages of music-syntactic processing, but they may have no significant impairment in late stages of implicit music-syntactic processing.

To our knowledge, this study is the first to demonstrate that individuals with congenital amusia exhibit morpho-syntactic deficits at an early stage of processing. Specifically, we found that amusics failed to display a classic ERP component – the so-called LAN – associated with an *early* stage of language morpho-syntactic processing (Friederici, 2002). In contrast, this ERP component was elicited normally in the control group. The absence of the ERAN and LAN in the amusia group supports the general hypothesis that music and language may share syntactic resources (Patel, 2003), especially at early stages of processing (Koelsch, 2012; Slevc and Okada, 2015). This finding complements a study on children with specific language impairment. Similar to our amusic participants, these children exhibited difficulties in processing morpho-syntactic and music-syntactic violations (Jentschke et al., 2008).

With regard to the later stage of syntactic processing, our findings suggest that amusics exhibited normal brain activity, as reflected by the presence of N5 and P600, which were also elicited in control participants by music-syntactic and morpho-syntactic incongruities, respectively. This outcome is consistent with previous studies (Koelsch et al., 2005; Steinbeis and Koelsch, 2008), which showed an interaction between music-syntactic and morpho-syntactic processing in normal adults, but only at the early stage of processing. Specifically, the LAN was significantly reduced when words were presented simultaneously with music-syntactic irregularities (Koelsch et al., 2005; Steinbeis and Koelsch, 2008). Furthermore, when comparing musicians and non-musicians, it was found that the ERAN amplitude was enhanced by musical training, but the N5 amplitude was not modulated by musical training (Jentschke and Koelsch, 2009; Koelsch et al., 2002b; Miranda and Ullman, 2007). This evidence suggests that the ERAN and N5 may reflect independent cognitive processes rather than a continuum of music-syntactic processing. Similarly, the LAN can be observed without the P600 as the two are dissociable (Mancini et al., 2011; Molinaro et al., 2015). Finally, our results confirm that the early language-processing deficit displayed by amusic participants is not a mere response to violation of any kind, but specific to syntax, as both amusia and control groups showed comparable N400 responses to semantic incongruities.

It should be noted that although we observed no impairment in late-stage music- and morpho-syntactic processing in individuals with congenital amusia, these findings do not necessarily imply that music and language syntactic processes are independent at a later stage. Rather, our findings indicate that the syntactic *impairments* exhibited by amusics are restricted to an early processing stage, and are unrelated to later stages of implicit syntactic processing. The ERAN and LAN belong to the

family of the MMN, which is thought to reflect any mismatch between top-down predictions and current inputs (Garrido et al., 2009), whereas P600 is considered to reflect the reanalysis, repair and integration of syntactic structure (Friederici, 2002). Following this line of reasoning, it appears that amusics lack the ability to properly predict the upcoming events (i.e., the final tones and words in the present study), as reflected by the absence of ERAN and LAN. However, when more time is given and a larger pattern of tones or words can be processed, amusics may be able to implicitly reanalyse and repair the anomalies to integrate the final tone/word into the whole melody/sentence, as indexed by the presence of N5 and P600.

Could the ERAN and LAN be caused by sensory violations instead of syntactic violations? Prior to our experiments, we ascertained that acoustic properties of the stimuli were adequately controlled. For the language task, plural nouns and singular nouns were included in both syntactically congruent and incongruent conditions. Thus, the mere presence or the absence of an “s” could not elicit any difference in ERP responses. For the music task, it could be argued that out of key notes result in both sensory and music-syntactic violations, making it difficult to interpret ERPs in response to such manipulations. However, there are compelling reasons to believe that the abnormal responses at early stages of processing in amusics reflect a music-syntactic impairment and not a sensory impairment. First, ERPs to sensory violations typically require an unexpected change to a constant attribute of sound, such as pitch or intensity. However, our syntactic violations occurred at a more abstract level, and acoustic attributes were not held constant prior to the violation. Second, previous studies have confirmed that, for typical listeners, an ERAN response is evoked by syntactic violations, even when sensory factors are taken into account (see Koelsch et al., 2007; Koelsch et al., 2002b; Omigie et al., 2013 but see Bigand et al., 2014). Third, there is already extensive evidence that bottom-up sensory information is successfully encoded in the primary auditory cortex of the amusic brain (Cousineau et al., 2015; Liu et al., 2014; Peretz, 2016). Therefore, if our manipulation had been processed as a sensory violation rather than a syntactic violation, then a normal brain response should have been observed in individuals with congenital amusia. However, our amusic participants exhibited abnormal brain responses to these violations, suggesting that they were processed as a syntactic violation, and not a sensory violation.

Neuroimaging studies on congenital amusia have identified structural and functional abnormalities within a fronto-temporal neural network (Albouy et al., 2013; Hyde et al., 2007; Hyde et al., 2006; Hyde et al., 2011; Loui et al., 2009; Mandell et al., 2007; but see Chen et al., 2015). This fronto-temporal network, in turn, is thought to contribute to music and language syntactic processing (Bianco et al., 2016; Janata et al., 2002; Koelsch et al., 2002c; Maess et al., 2001; Sammler et al., 2013; Tillmann et al., 2006). In particular, temporal generators, as well as prefrontal generators, may contribute to ERAN and LAN (Hanna et al., 2014; Maess et al., 2001). Collectively, all these findings help to account for the present finding that the impairments associated with congenital amusia not only lead to music-syntactic deficits, but also lead to impaired language-syntactic processing at a neurological level. Whether amusic individuals also exhibit subtle deficits for language syntax in their daily life is currently unknown, and awaits future investigation.

Conflict of interest

The authors declare no competing financial interests.

Funding

This research was supported by the Australian Research Council Discovery Grant (DP130101084), awarded to W.F.T.

References

- Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P.E., Daligault, S., Delpuech, C., Bertrand, O., Caclin, A., Tillmann, B., 2013. Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex. *Brain* 136, 1639–1661.
- Asano, R., Boeckx, C., 2015. Syntax in language and music: what is the right level of comparison? *Front. Psychol.* 6.
- Ayotte, J., Peretz, I., Hyde, K., 2002. Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain* 125, 238–251.
- Besson, M., Faita, F., 1995. An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1278.
- Besson, M., Macar, F., 1987. An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology* 24, 14–25.
- Bharucha, J.J., 1984. Anchoring effects in music: the resolution of dissonance. *Cogn. Psychol.* 16, 485–518.
- Bianco, R., Novembre, G., Keller, P.E., Kim, S.G., Scharf, F., Friederici, A.D., Villringer, A., Sammler, D., 2016. Neural networks for harmonic structure in music perception and action. *NeuroImage* 142, 454–464.
- Bigand, E., Tillmann, B., Poulin-Charronnat, B., 2006. A module for syntactic processing in music? *Trends Cogn. Sci.* 10, 195–196.
- Bigand, E., Delbé, C., Poulin-Charronnat, B., Leman, M., Tillmann, B., 2014. Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. *Front. Syst. Neurosci.* 8.
- Boersma, P., Weenink, D., 2014. Praat: Doing Phonetics by Computer.
- Braun, A., Mcardle, J., Jones, J., Nechaev, V., Zaleski, C., Brewer, C., Drayna, D., 2008. Tune deafness: processing melodic errors outside of conscious awareness as reflected by components of the auditory ERP. *PLoS One* 3, e2349.
- Carrus, E., Pearce, M.T., Bhattacharya, J., 2013. Melodic pitch expectation interacts with neural responses to syntactic but not semantic violations. *Cortex* 49, 2186–2200.
- Chen, J.L., Kumar, S., Williamson, V.J., Scholz, J., Griffiths, T.D., Stewart, L., 2015. Detection of the arcuate fasciculus in congenital amusia depends on the tractography algorithm. *Front. Psychol.* 6, 9.
- Cousineau, M., McDermott, J.H., Peretz, I., 2012. The basis of musical consonance as revealed by congenital amusia. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19858–19863.
- Cousineau, M., Oxenham, A.J., Peretz, I., 2015. Congenital amusia: a cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia* 66, 293–301.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21.
- Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., Gibson, E., 2009. Structural integration in language and music: evidence for a shared system. *Mem. Cogn.* 37, 1–9.
- Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84.
- Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463.
- Gouvea, A.C., Phillips, C., Kazanina, N., Poeppel, D., 2010. The linguistic processes underlying the P600. *Lang. Cogn. Process.* 25, 149–188.
- Hanna, J., Mejias, S., Schelstraete, M.A., Pulvermuller, F., Shtyrov, Y., Van der Lely, H.K., 2014. Early activation of Broca's area in grammar processing as revealed by the syntactic mismatch negativity and distributed source analysis. *Cogn. Neurosci.* 5, 66–76.
- Hasting, A.S., Kotz, S.A., Friederici, A.D., 2007. Setting the stage for automatic syntax processing: the mismatch negativity as an indicator of syntactic priming. *J. Cogn. Neurosci.* 19, 386–400.
- Henry, M.J., McAuley, J.D., 2013. Failure to apply signal detection theory to the Montreal Battery of Evaluation of Amusia may misdiagnose amusia. *Music. Percept.* 30, 480–496.
- 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, 2562–2570.
- 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, 13028–13032.
- Hyde, K.L., Zatorre, R.J., Peretz, I., 2011. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cereb. Cortex* 21, 292–299.
- Janata, P., Birk, J.L., Van Horn, J.D., Leman, M., Tillmann, B., Bharucha, J.J., 2002. The cortical topography of tonal structures underlying Western music. *Science* 298, 2167–2170.
- Jentschke, S., Koelsch, S., 2009. Musical training modulates the development of syntax processing in children. *NeuroImage* 47, 735–744.
- Jentschke, S., Koelsch, S., Sallat, S., Friederici, A.D., 2008. Children with specific language impairment also show impairment of music-syntactic processing. *J. Cogn. Neurosci.* 20, 1940–1951.
- Jiang, C., Liu, F., Thompson, W.F., 2016. Impaired explicit processing of musical syntax and tonality in a group of mandarin-speaking congenital amusics. *Music. Percept.* 33, 401–413.
- Koelsch, S., 2005. Neural substrates of processing syntax and semantics in music. *Curr. Opin. Neurobiol.* 15, 207–212.
- Koelsch, S., 2011. Toward a neural basis of music perception – a review and updated model. *Front. Psychol.* 2, 110.
- Koelsch, S., 2012. *Brain and Music*. Wiley-Blackwell, West Sussex.
- Koelsch, S., Friederici, A.D., 2003. Toward the neural basis of processing structure in music. Comparative results of different neurophysiological investigation methods. *Ann. N. Y. Acad. Sci.* 999, 15–28.
- Koelsch, S., Jentschke, S., 2010. Differences in electric brain responses to melodies and chords. *J. Cogn. Neurosci.* 22, 2251–2262.
- Koelsch, S., Gunter, T., Friederici, A.D., Schroger, E., 2000. Brain indices of music processing: “nonmusicians” are musical. *J. Cogn. Neurosci.* 12, 520–541.
- Koelsch, S., Schroger, E., Gunter, T.C., 2002a. Music matters: preattentive musicality of the human brain. *Psychophysiology* 39, 38–48.
- Koelsch, S., Schmidt, B.H., Kansok, J., 2002b. Effects of musical expertise on the early right anterior negativity: an event-related brain potential study. *Psychophysiology* 39, 657–663.
- Koelsch, S., Gunter, T.C., v Cramon, D.Y., Zysset, S., Lohmann, G., Friederici, A.D., 2002c. Bach speaks: a cortical “language-network” serves the processing of music. *NeuroImage* 17, 956–966.
- Koelsch, S., Gunter, T.C., Wittfoth, M., Sammler, D., 2005. Interaction between syntax processing in language and in music: an ERP Study. *J. Cogn. Neurosci.* 17, 1565–1577.
- Koelsch, S., Jentschke, S., Sammler, D., Mietschen, D., 2007. Untangling syntactic and sensory processing: an ERP study of music perception. *Psychophysiology* 44, 476–490.
- Kunert, R., Willems, R.M., Casasanto, D., Patel, A.D., Hagoort, P., 2015. Music and language syntax interact in Broca's area: an fMRI study. *PLoS One* 10, e0141069.
- Kunert, R., Willems, R.M., Hagoort, P., 2016. Language influences music harmony perception: effects of shared syntactic integration resources beyond attention. *R. Soc. Open Sci.* 3, 150685.
- Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470.
- Liu, F., Patel, A.D., Fourcin, A., Stewart, L., 2010. Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain* 133, 1682–1693.
- Liu, F., Maggu, A.R., Lau, J.C., Wong, P.C., 2014. Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers. *Front. Hum. Neurosci.* 8, 1029.
- Loui, P., Alsop, D., Schlaug, G., 2009. Tone deafness: a new disconnection syndrome? *J. Neurosci.* 29, 10215–10220.
- Maess, B., Koelsch, S., Gunter, T.C., Friederici, A.D., 2001. Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.* 4, 540–545.
- Mancini, S., Molinaro, N., Rizzi, L., Carreiras, M., 2011. When persons disagree: an ERP study of Unagreement in Spanish. *Psychophysiology* 48, 1361–1371.
- Mandell, J., Schulze, K., Schlaug, G., 2007. Congenital amusia: an auditory-motor feedback disorder? *Restor. Neurol. Neurosci.* 25, 323–334.
- Mignault Goulet, G., Moreau, P., Robitaille, N., Peretz, I., 2012. Congenital amusia persists in the developing brain after daily music listening. *PLoS One* 7, e36860.
- Miranda, R.A., Ullman, M.T., 2007. Double dissociation between rules and memory in music: an event-related potential study. *NeuroImage* 38, 331–345.
- Molinaro, N., Barber, H.A., Caffarra, S., Carreiras, M., 2015. On the left anterior negativity (LAN): the case of morphosyntactic agreement: a reply to Tanner et al. *Cortex* 66, 156–159.
- Moreau, P., Jolicoeur, P., Peretz, I., 2013. Pitch discrimination without awareness in congenital amusia: evidence from event-related potentials. *Brain Cogn.* 81, 337–344.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K., 2007. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Omigie, D., Pearce, M.T., Williamson, V.J., Stewart, L., 2013. Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia* 51, 1749–1762.
- Paller, K.A., McCarthy, G., Wood, C.C., 1992. Event-related potentials elicited by deviant endings to melodies. *Psychophysiology* 29, 202–206.
- Patel, A.D., 2003. Language, music, syntax and the brain. *Nat. Neurosci.* 6, 674–681.
- Patel, A.D., Gibson, E., Ratner, J., Besson, M., Holcomb, P.J., 1998. Processing syntactic relations in language and music: an event-related potential study. *J. Cogn. Neurosci.* 10, 717–733.
- Patel, A.D., Iversen, J.R., Wassenaar, M., Hagoort, P., 2008. Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology* 22, 776–789.
- Peretz, I., 1993. Auditory ataxia for melodies. *Cogn* 10, 21–56.
- Peretz, I., 2016. Neurobiology of congenital amusia. *Trends Cogn. Sci.* 20, 857–867.
- Peretz, I., Ayotte, J., Zatorre, R.J., Mehler, J., Ahad, P., Penhune, V.B., Jutras, B., 2002. Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron* 33, 185–191.
- Peretz, I., Champod, A.S., Hyde, K., 2003. Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Ann. N. Y. Acad. Sci.* 999, 58–75.
- Peretz, I., Cummings, S., Dube, M.P., 2007. The genetics of congenital amusia (tone deafness): a family-aggregation study. *Am. J. Hum. Genet.* 81, 582–588.
- Peretz, I., Gosselin, N., Tillmann, B., Cuddy, L.L., Gagnon, B., Trimmer, C.G., Paquette, S., Bouchard, B., 2008. On-line identification of congenital amusia. *Music. Percept.* 25, 331–343.
- Peretz, I., Brattico, E., Järvenpää, M., Tervaniemi, M., 2009. The amusic brain: in tune, out of key, and unaware. *Brain* 132, 1277–1286.
- Peretz, I., Vuvan, D., Lagrois, M., Armony, J.L., 2015. Neural overlap in processing music and speech. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 370.
- Peter, V., McArthur, G., Thompson, W.F., 2010. Effect of deviance direction and calculation method on duration and frequency mismatch negativity (MMN). *Neurosci. Lett.* 482, 71–75.
- Pfeifer, J., Hamann, S., 2015. Revising the diagnosis of congenital amusia with the Montreal Battery of Evaluation of Amusia. *Front. Hum. Neurosci.* 9, 161.
- Sammler, D., Koelsch, S., Friederici, A.D., 2011. Are left fronto-temporal brain areas a

- prerequisite for normal music-syntactic processing? *Cortex* 47, 659–673.
- Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H.J., Knosche, T.R., Wellmer, J., Widman, G., Elger, C.E., Friederici, A.D., Schulze-Bonhage, A., 2013. Co-localizing linguistic and musical syntax with intracranial EEG. *NeuroImage* 64, 134–146.
- Sarkamo, T., Tervaniemi, M., Soinila, S., Autti, T., Silvennoinen, H.M., Laine, M., Hietanen, M., 2009. Cognitive deficits associated with acquired amusia after stroke: a neuropsychological follow-up study. *Neuropsychologia* 47, 2642–2651.
- Slevc, L.R., Okada, B.M., 2015. Processing structure in language and music: a case for shared reliance on cognitive control. *Psychon. Bull. Rev.* 22, 637–652.
- Slevc, L.R., Rosenberg, J.C., Patel, A.D., 2009. Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychon. Bull. Rev.* 16, 374–381.
- Slevc, L.R., Farooqi-Shah, Y., Saxena, S., Okada, B.M., 2016. Preserved processing of musical structure in a person with agrammatic aphasia. *Neurocase* 22, 505–511.
- Stanislaw, H., Todorov, N., 1999. Calculation of signal detection theory measures. *Behav. Res. Methods Instrum. Comput.* 31, 137–149.
- Steinbeis, N., Koelsch, S., 2008. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cereb. Cortex* 18, 1169–1178.
- Steinhauer, K., Drury, J.E., 2012. On the early left-anterior negativity (ELAN) in syntax studies. *Brain Lang.* 120, 135–162.
- Sun, Y., Lu, X., Ho, H.T., Thompson, W.F., 2017. Pitch discrimination associated with phonological awareness: evidence from congenital amusia. *Sci. Rep.* 7, 44285.
- Thompson, W.F., 1993. Modeling perceived relationships between melody, harmony, and key. *Percept. Psychophys.* 53, 13–24.
- Thompson, W.F., Cuddy, L.L., 1989. Sensitivity to key change in chorale sequences: a comparison of single voices and four-voice harmony. *Music. Percept.* 7, 151–168.
- Thompson, W.F., Cuddy, L.L., 1992. Perceived key movement in four-voice harmony and single voices. *Music. Percept.* 9, 427–438.
- Thompson, W.F., Marin, M.M., Stewart, L., 2012. Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 109, 19027–19032.
- Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A.D., von Cramon, D.Y., 2006. Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *NeuroImage* 31, 1771–1782.
- Tillmann, B., Gosselin, N., Bigand, E., Peretz, I., 2012. Priming paradigm reveals harmonic structure processing in congenital amusia. *Cortex* 48, 1073–1078.
- Van de Cavey, J., Hartsuiker, R.J., 2016. Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition* 146, 172–184.
- Widmann, A., Schroger, E., 2012. Filter effects and filter artifacts in the analysis of electrophysiological data. *Front. Psychol.* 3, 233.
- Zendel, B.R., Lagrois, M., Robitaille, N., Peretz, I., 2015. Attending to pitch information inhibits processing of pitch information: the curious case of amusia. *J. Neurosci.* 35, 3815–3824.