Neuropsychologia 51 (2013) 1749–1762

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

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

Electrophysiological correlates of melodic processing in congenital amusia

Diana Omigie [a,*], Marcus T. Pearce [b], Victoria J. Williamson [a], Lauren Stewart [a]

[a] Goldsmiths, University of London, London SE14 6NW, UK
[b] Queen Mary, University of London, London E1 4NS, UK

ARTICLE INFO

Article history:
Received 18 July 2012
Received in revised form
25 April 2013
Accepted 10 May 2013
Available online 23 May 2013

Keywords:
Congenital amusia
Melodic processing
Electroencephalography
Expectations
Explicit knowledge

ABSTRACT

Music listening involves using previously internalized regularities to process incoming musical structures. A condition known as congenital amusia is characterized by musical difficulties, notably in the detection of gross musical violations. However, there has been increasing evidence that individuals with the disorder show preserved musical ability when probed using implicit methods. To further characterize the degree to which amusic individuals show evidence of latent sensitivity to musical structure, particularly in the context of stimuli that are ecologically valid, electrophysiological recordings were taken from a sample of amusic and control participants as they listened to real melodies. To encourage them to pay attention to the music, participants were asked to detect occasional notes in a different timbre. Using a computational model of auditory expectation to identify points of varying levels of expectedness in these melodies (in units of information content (IC), a measure which has an inverse relationship with probability), ERP analysis investigated the extent to which the amusic brain differs from that of controls when processing notes of high IC (low probability) as compared to low IC ones (high probability). The data revealed a novel effect that was highly comparable in both groups: Notes with high IC reliably elicited a delayed P2 component relative to notes with low IC, suggesting that amusic individuals, like controls, found these notes more difficult to evaluate. However, notes with high IC were also characterized by an early frontal negativity in controls that was attenuated in amusic individuals. A correlation of this early negative effect with the ability to make accurate note expectedness judgments (previous data collected from a subset of the current sample) was shown to be present in typical individuals but compromised in individuals with amusia: a finding in line with evidence of a close relationship between the amplitude of such a response and explicit knowledge of musical deviance.

1. Introduction

While most individuals show a natural aptitude for music perception and production, individuals with a condition known as congenital amusia (amusia, hereafter) show considerable problems in this regard. The disorder is believed to affect 4% of the general population based on a study where a large sample of the British population were assessed on their ability to detect pitch errors in popular tunes (Kalmus & Fry, 1980: Although see Henry & McAuley (2010), for a critique of the way this statistic is reached). Since then, a similar estimate has been obtained using the current prevailing tool for diagnosing the disorder: the Montreal Battery for the Evaluation of Amusia (MBEA) (Peretz, Champod, & Hyde, 2003). Encompassing a range of subtests that assess various aspects of musical processing, the MBEA has associated amusia with severe impairments along the pitch dimension of musical processing. Specifically, affected individuals show difficulties in recognizing changes in intervallic structure and detecting out of key notes in the context of a melody (Ayotte, Peretz, & Hyde, 2002; Peretz, Ayotte, Zatorre, Mehler, Pehune, & Jutras, 2002; Peretz et al., 2003).

Studies testing twin pairs and first-degree relatives of individuals with amusia suggest a genetic basis for the condition (Drayna, Manichaikul, de Lange, Snieder, & Spector, 2001; Peretz, Cummings, & Dubé, 2007) and structural neuro-imaging studies associate the condition with subtle neurological abnormalities. In particular, differences in brain structure have been reported in inferior frontal cortex and superior temporal areas in both the left and right hemisphere (Hyde et al., 2007; Hyde, Zatorre, Griffiths, Lerch, & Peretz, 2006; Mandell, Schulze, & Schlaug, 2007). Further, results from a study using diffusion tensor imaging suggest that individuals with amusia have reduced structural connectivity in the right superior branch of the *arcuate fasciculus*, a large fibre bundle connecting temporal and frontal areas of the brain (Loui,

* Corresponding author. Tel.: +44 207 078 5465.
E-mail address: d.omigie@gmail.com (D. Omigie).

0028-3932/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2013.05.010

Wu, Wessel, & Knight, 2009a). In terms of functional studies of amusia, an fMRI study (Hyde et al., 2007) was able to confirm the role of areas previously implicated by structural imaging studies (Hyde et al., 2006). Specifically, this study revealed a global functional brain difference between controls and amusics in response to changing pitch sequences, whereby amusics showed both reduced connectivity of the pars orbitalis of the right IFG with auditory cortex, as well as increased connectivity between the right and left auditory cortex. In contrast, there was no difference in the extent to which both the amusic and control auditory areas showed a positive linear increase in blood–oxygen-level-dependent (BOLD) response as a function of increasing pitch distance between successive tones.

Event-related potentials (ERPs) are especially useful for probing the functional aspects of congenital amusia, since the high temporal resolution of this approach is ideally suited to measuring neural processing of dynamically evolving stimuli, such as music. Several previous studies have used the ERP approach to investigate how the amusic brain processes musical or pitch-related information (Moreau, Jolicoeur, & Peretz, 2009; Peretz, Brattico, & Tervaniemi, 2005; Peretz, Brattico, Järvenpää, & Tervaniemi, 2009). Overall, these studies have raised the interesting possibility that the brains of individuals with amusia process aspects of pitch that they are unable to report (Moreau et al., 2009; Peretz, et al., 2009). However, the use of simple oddball stimuli and manipulated melodies limit the extent to which these studies' findings can be generalized to the processes involved in everyday music listening. The current study aimed to investigate melodic pitch processing in amusia, examining the extent to which such individuals show evidence of spared processing of musical structure in the context of ecologically valid stimuli.

When considering typical individuals, it is widely acknowledged that expectancy, described as the anticipation of an event based on its probability of occurring (Chaplin, 1985), plays an important role, not just in the aesthetic and emotional aspects of musical listening (Huron, 2006; Juslin & Vastfjall, 2008) but also in explaining how listeners recognize and remember music (Schmuckler, 1997; Schulkind, Posner, & Rubin, 2003). One particularly influential account of the source of musical expectations is that listeners internalize the patterns of occurrence and co-occurrence of musical events in music to acquire a sophisticated knowledge of musical structure over a lifetime of listening (Tillmann, Bharucha, & Bigand, 2000). In turn, this account of the source of musical expectations has inspired a computational model of melodic expectation, based on information theory and statistical learning (Pearce, 2005; Pearce & Wiggins, 2006; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010).

The basis of the model's predictions is that the cognitive processes that generate expectations weigh the probability of different possible continuations to a musical excerpt based on the frequency with which different continuations have followed similar contexts in their previous experience (Pearce, 2005). In the model, the expectedness of the individual notes in a melody is expressed in units of Information content (IC), where IC (the negative logarithm, to the base 2, of the probability of an event occurring) is a lower bound on the number of bits required to encode an event in context (Mackay, 2003). In short, low IC notes are the notes that have high probability in the context of a melody and should therefore be 'expected' to a listener while high IC notes are those that have a low probability in the context of the melody and should therefore be 'unexpected'. In line with previous papers using the model, we henceforth describe musical events in terms of their IC.

Based as it is, on the notion that melodic expectations arise solely from statistical learning, the Pearce (2005) model is arguably more parsimonious than previous approaches to modeling melodic expectation. Perhaps the most influential account of melodic expectations came from Narmour (1990) who suggested that listeners' expectations are influenced by two independent cognitive systems: *bottom up* influences which comprise innate and universal gestalt-like principles, and *style specific* influences, which develop through continued exposure to a given musical style. Narmour's implication–realization model found support in a series of experimental studies that examined the bottom up principles he outlined (e.g. Cuddy & Lunny, 1995; Krumhansl & Kessler, 1982). However after carrying out an independent analysis of the data, Schellenberg (1997) argued that bottom up models proposed by Narmour and Krumhansl are overspecified and may be expressed more parsimoniously.

Schellenberg's model, which suggested that two factors, namely *principle of proximity* (consecutive notes tend to be proximate in pitch) and *pitch reversal* (a tendency for registral direction change), are sufficient to explain listeners' expectation, did indeed show greater simplicity along with comparable predictive power. However, it was necessarily limited in making only local pitch predictions based on the preceding one or two notes. In contrast, the model of Pearce (2005) predicts which pitches will occur based on preceding melodic contexts of varying lengths. Critically, it has been shown to outperform Schellenberg's two-factor model in predicting listeners' subjective expectations (Pearce, 2005; Pearce & Wiggins, 2006; Pearce et al., 2010) with results from multiple regression analyses revealing that it accounted for more variance in the ratings and response times of a group of typical listeners than the two-factor model (78% of the variance in the ratings and 56% of the variance in the response times compared to approximately 56% and 33%, respectively) (Pearce et al., 2010).

Another important property of the model, which arguably makes it a superior choice for modelling melodic expectations, is its use of a long-term and a short-term component to simulate how expectations are formed when a given piece of music is presented. The long-term model component is trained on a corpus of western tonal melody, which represents schematic expectations learned over a lifetime of exposure while the short-term model is trained incrementally for each melody that it is presented with, to simulate local influences on expectations that are formed dynamically as a given piece of music unfolds. The building of predictions based on learned regularities on both of these time scales provides a degree of sophistication that is lacking in other models.

In a previous behavioural study (Omigie, Pearce, & Stewart, 2012), we examined how amusic listeners responded to notes of high or low IC based on the predictions of this computational model of melodic expectation. In that study, two versions of a melodic priming paradigm were used to probe participants' abilities to form melodic pitch expectations, in an implicit and an explicit manner. In the implicit version, participants made speeded, forced-choice discriminations concerning the timbre of a cued target note. In the explicit version, participants used a 1–7 rating scale to indicate the degree to which the pitch of the cued target note was expected or unexpected. We observed that amusics and controls were no different in the extent to which they showed evidence of implicit musical expectations but that amusics were significantly worse than controls at using explicit ratings to differentiate between high and low IC events in a melodic context. In another recent study, also using a priming paradigm, but this time, to investigate processing of harmonic structure, amusic participants were shown to be facilitated in their processing of functionally important as opposed to less important chords in the context of chord sequences, providing further evidence that amusic participants can develop expectancies for musical events at an implicit level (Tillmann, Gosselin, Bigand, & Peretz, 2012). The current study sought to further investigate this discrepancy between the implicit and explicit music anticipatory capacities of those with congenital amusia by collecting electrophysiological recordings from a sample of such individuals and control participants as they listened to real melodies.

In typical listeners, violations of musical expectations have been associated with a number of ERP components but one in particular has received a great deal of attention due to its presence even when no task is required of the listener. This early negative response occurring at around 150 ms post-onset of the deviant musical event has been termed the 'Early right anterior negativity' or ERAN (Koelsch, Gunter, Friederici, & Schröger, 2000; Koelsch, Schmidt & Kansok, 2002; Leino, Brattico, Tervaniemi, & Vuust, 2007) although it is sometimes also referred to as the 'Early anterior negativity' when no lateralization is observed (Koelsch, Schröger, & Tervaniemi, 1999; Loui, Grent-'t-Jong, Torpey, & Woldorff, 2005). The ERAN may be considered as the musical syntactic version of the Mismatch Negativity, MMN, an ERP component of similar latency and topography (Näätänen, Paavilainen, Rinne, & Alho, 2007).

The ERAN and MMN are often distinguished based on the fact that the MMN is elicited in response to regularities internalized online, during the listening session, while the ERAN is elicited in response to violations of rules present in long-term musical knowledge. However, they are both similar in being elicited by deviant events that have a low probability of occurring in an auditory stream (i.e. high IC). In the case of the MMN, this is in relation to an ongoing stream of standard events while in the case of the ERAN, this is in relation to the local context as opposed to the overall probability of the event occurring. Nevertheless, it has been suggested that the two kinds of neural signature may be based on the same mechanism of probabilistic learning. Loui et al. (2009b) showed that the time course and scalp topographies of the ERP response to violations within an artificially constructed music system are identical to those observed when violations are encountered while listening to stylistically familiar music, while Kim, Kim, and Chung (2011) showed that neuro-magnetic responses to musical chords correlate with the probability of that chord occurring in a representative sample of Western tonal music. Importantly, while the ERAN is typically associated with harmonic violations, several studies have also reported a similar early negative response, at the latency of the N1, to violations in the context of monophonic melodies (Koelsch & Jentschke, 2010; Miranda & Ullman, 2007; Loui et al., 2009b).

Based on the evidence that melodic violations result in a negative deflection at the latency of the N1 (Koelsch & Jentschke, 2010), we examined the amplitude and latency of this component in the current study. While the current stimuli used were all ecologically valid real melodies, we assumed that the processing of notes of high IC, even in the context of real melodies, would be underpinned by the same processes that are involved in the processing of gross musical violations such as out of key notes. As the size of the early negative response elicited in a musical context (the ERAN) has been shown to be related to the probability of an event occurring (Kim et al., 2011; Loui et al., 2009b) we predicted that the size of the observed early negative response in controls would correlate with the degree of note expectedness as predicted by our model. However, as the early negative response has also been shown to correlate with conscious awareness of a musical event as a deviant (Koelsch et al., 1999; Koelsch et al., 2002; Koelsch, Jentschke, Sammler, & Mietchen, 2007; Miranda & Ullman, 2007) we predicted that individuals with amusia – who show impaired sensitivity to gross musical violations at a behavioural level – might, relative to controls, show an attenuated early negative response to low probability notes in real melodies. In addition, as the influence of tonal expectations has been shown on a number of other ERP components, even as early as within the first 100 ms after tone onset (e.g. Marmel, Perrin & Tillmann, 2011), we systematically examined other obligatory components of the auditory evoked potential, the P1 and P2, to investigate whether there is any effect of note IC on the amplitude and latency of these responses (Naatanen, 1992).

In the first instance, we carried out two types of ERP analyses to examine the effect of IC on components of the auditory evoked potential. The first of these, the primary analysis, was comparable to the traditional approach taken when investigating ERP correlates of musical violations, whereby responses to 'regular' and 'irregular' events are compared in order to identify electrophysiological correlates of musical expectation. In the present study, this involved the selection of notes of low, medium and high IC using our computational model and comparing the mean amplitude and latency of the obligatory ERP responses across these event categories. The second type, the secondary analysis, sought to characterize any categorical effects identified in the primary analysis in greater detail, specifically verifying whether a linear relationship could be observed between note IC and these effects. Accordingly, in the secondary analysis, the notes of each melody were sorted by their IC and assigned to ten categories of increasing IC thus allowing the parametric relationship between IC and any observed ERP effects to be further examined using correlation analyses.

In addition to these two initial analyses carried out to examine the effect of IC on components of the auditory evoked potential, we carried out further supporting and exploratory analyses to different ends. The first of these analyses aimed to clarify the behavioural relevance of the observed ERP effects by relating them to previously collected behavioural data (Omigie et al., 2012), while the second aimed to assess whether relationship between any observed ERP effects and model IC predictions was modulated over time. Finally, the third examined both the extent to which the two groups showed evidence of adapting to the mean IC within a given melody (which may vary across melodies) as well as the possibility that they responded differently to the different musical features and components of the model configuration used in the main analyses.

2. Methods

2.1. Participants

A total of 30 participants (15 amusic, 15 control) took part in the study. All participants were recruited via an online assessment based on the scale and rhythm subtests of the MBEA (Peretz et al., 2003) (www.delosis.com/listening/home.html). Each participant took the online test twice and those who consistently achieved a score of 22/30 or below were invited to come to the laboratory where assessment could take place under controlled conditions. Four MBEA subtests (scale, contour, interval and rhythm subtests) were administered in a sound attenuated booth in order to confirm the presence or absence of amusia. Previous research has shown that amusia is characterized by poor perception in the pitch-based subtests of the MBEA (scale, contour, interval) while only half of the individuals with amusia typically show a deficit in the rhythm test (Peretz et al., 2003). Thus we calculated a composite score for the three pitch-based subtests, using 65 as a cut off score, whereby individuals were classified as amusic if their composite score fell below this value (Liu, Patel, Fourcin, & Stewart, 2010; Williamson & Stewart, 2010). Table 1 provides background information on the two groups in terms of age, gender, number of years of formal education and number of years of musical education. Table 2 provides scores on the MBEA subtests and psychophysically measured pitch change detection and pitch direction discrimination thresholds that we include as an additional background measure (see Liu et al., 2010; Williamson, Liu, Peryer, Grierson, & Stewart, 2012, for procedural details).

Table 1
Descriptive statistics and results of t-tests comparing amusic and control participant characteristics.

		Age	Gender	Years of Ćmusical training	Years of Ćeducation*
Amusic	Mean	56.27	10 F	0.27	16.23
	SD	8.51	5 M	1.03	1.96
Control	Mean	50.53	10 F	0.75	16.3
	SD	10.74	5 M	1.62	2.46
t-tests	*t*	1.62		1.00	0.09
	df	*28*		28	24
	p	.12		.34	.93

* Years of education: missing data from two amusic participants. mean, SD and t -tests computed using reduced sizes for both groups.

Table 2
Descriptive statistics and results of *t*-tests comparing performance of amusic and control participants on subtests of the Montreal Battery of Evaluation of Amusia (MBEA) and psychophysically measured pitch thresholds.

		MBEA Ćscale	MBEA Ćcontour	MBEA Ćinterval	MBEA Ćrhythm	Pitch Ćcomposite	Detection* ĆThreshold (semitones)	Detection* ĆThreshold (semitones)
Amusic	Mean	19.4	19.73	18.27	23.67	56.67	0.29	1.68
	SD	2.22	2.55	1.62	3.5	5.19	0.3	1.38
Control	Mean	27.67	27.93	28.00	28.27	83.6	0.14	0.18
	SD	1.63	2.15	2.20	1.39	5.14	0.06	0.10
t-tests	*T*	11.58	9.51	13.77	5.47	14.0	1.44	3.58
	df		28	28	28	28	20	20
	p	28 < .001	< .001	< .001	< .001	< .001	.18	.005

* Detection and direction thresholds: missing data from one amusic and four control participants. Mean, SD and *t* -tests computed using reduced sizes for both groups.

2.2. Stimuli

2.2.1. Musical material

The stimuli consisted of 58 hymns (including 2 practice trials), randomly selected and transcribed from a Church of England hymnal (Nicholson, Knight, Dykes, & Bower, 1950). The melodies were in a range of major keys (3 in A flat, 17 in E flat, 5 in B flat, 8 in F, 4 in C, 10 in G, 5 in D, 3 in A and 3 in E major keys, respectively). Individual notes were created using the electronic piano 1 instrument of a Roland sound canvas (SC-88) MIDI synthesizer before being converted to wav files. In order to focus on pitch expectations in particular, the rhythmic structure of the melodies was removed in a musically sensitive manner by a skilled musicologist so that each note had the same duration of 600 ms and an equivalent inter-onset interval of 700 ms, with the amplitude of each note kept constant. The melodies varied in length from 32 to 64 notes reflecting the range of melody lengths present in the hymnal (47 melodies of 32 notes length, 9 melodies of 48 notes length and 2 melodies of 64 notes length) and were the same as those used in Experiment 2 (the implicit task) of our previous study (Omigie et al., 2012). Individual sound files for each note were presented using an E-Prime program, which played each melody in turn. In 6 out of the 56 melodies presented in the experiment, a single note was modified to play in a different timbre (the electric grand piano instrument of the Roland sound canvas (SC-88) MIDI synthesizer). Sample stimuli may be seen on the journal website. We did not consider metre here because these are simple melodies played without dynamic accents whose metrical interpretation is relatively unambiguous and previous research with these melodies has demonstrated that the model is capable of accounting well for listener's expectations without explicitly modeling any potential effects of meter (Pearce et al., 2010; Omigie et al., 2012). However, modeling the effects of metrical interpretation on pitch expectation remains an important topic for future research with other stimuli.

2.2.2. Selecting the probe notes

As explained above, points of varying expectedness in each melody were objectively defined using a computational model of melodic expectation which is based on the theory that listeners weigh the probability of different possible continuations to a musical excerpt based on the frequency with which different continuations have followed a similar context in their previous experience (Huron 2006; Pearce, 2005; Pearce & Wiggins, 2006; Pearce et al., 2010).

The predictions of the model may be based solely on a long-term model (LTM), trained on a corpus of melodies (western tonal melodies in this case) and representing schematic expectations learned over a lifetime of exposure or solely on a short-term model (STM) trained incrementally for each melody and simulating the ongoing listening experience whereby expectations are formed dynamically as the music unfolds, or, in the most realistic case which has been shown to best account for participant data, on a combination of both. In this latter case, the predictions of the long- and short-term models are united by assigning greater weights to the model whose predictions are associated with lower uncertainty, to produce a single probability distribution. Specifically, the combination process involves taking the product of the weighted probability estimates returned by each model for each possible value and normalizing it such that combined estimates sum to one over the possible pitches that could occur (the pitch alphabet). The improved performance in predicting novel melodies that this approach allows, relative to other model combination methods, is described in greater detail in previous work (e.g. Pearce, 2005). In the main analyses in this study (the primary and secondary analyses), the model computed probabilistic pitch predictions using a combined representation of the given note's scale degree, relative to the tonic of the notated key of the melody, as well as the size and direction of the interval preceding it, as this has been shown to be a good predictor of listeners' expectations. A combination of the long- and short-term model was used for the same reason.

Probe notes were selected in different ways for the two types of analysis (primary and secondary). In the primary analysis, designed to observe which ERP components showed sensitivity to note expectedness, two notes were selected from the low, medium and high range of the IC profile of each melody (based on the ScaleDeg*Int, combined LTM & STM version of the model). In the secondary analysis, carried out to further explore the relationship between observed ERP effects and IC, all but the first two notes in each melody[1] were combined to make one large set. The members of this set, containing all the notes to be analysed, were then sorted in order of increasing IC irrespective of melody membership and the 10% of notes with the lowest IC were assigned to bin 1, the next 10% of notes with the lowest IC assigned to bin 2, and so on and so forth. Fig. 1A shows the IC profile of a sample melody used in the experiment, Fig. 1B shows the distribution of IC of probe notes in the primary analysis superimposed on the distribution of all the notes in the 58 hymns and Fig. 1C shows the mean IC of each of the 10 bin categories in the secondary analysis. Table 3 shows properties of these probe notes in the primary and secondary analysis including the mean tonal stability values (Krumhansl & Kessler, 1982), the size of the preceding interval[2] as well as the mean position of the notes within the melody.

2.3. Procedure

Participants were seated in front of a computer monitor in a dark, quiet testing room. The stimuli (58 melodies) were each presented once at a comfortable listening volume through speakers placed behind the participant. The stimuli were presented using the software E-prime in three blocks (two comprising 20 melodies and one comprising 18) lasting approximately 12 min each. The melodies in each block were presented in randomized order. Two melodies in each block contained a deviant timbre, such that six deviants in total, across the three blocks, had to be detected. Participants were instructed to listen to each melody with their eyes closed and detect whether any note in the melody had been played in a different timbre. They were asked to indicate, using a response box, whether or not they had heard a change in timbre. Responses were given after a melody was heard. The purpose of this task was to ensure that participants attended to the stimuli during the EEG recording session. Two practice trials, both of which contained the target timbre, were presented to familiarize the participants with the procedure.

2.4. EEG recording

Participants' EEG was measured using the Neuroscan measuring system (Neuroscan SynAmps2; Compumedics, El Paso, TX). Scalp EEG was recorded at a sampling rate of 500 Hz, using 64 electrodes mounted into an elastic cap. Bipolar vertical and horizontal electro-oculograms (EOG) were recorded from four additional channels to monitor eye movements and blinks. Electrode impedances were kept below 5 kΩ. The average of two ear electrodes (one from each earlobe) was used as a reference. Preprocessing of the raw data was carried out using batch scripts created with the EEGLAB toolbox (Delorme & Makeig, 2004) for MATLAB (The Mathworks Inc, Natick). Raw EEG data were subjected to a low pass filter of 70 Hz and a notch filter (45–55 Hz) was applied to remove power line noise. Data epochs representing single trials time-locked to the onset of the target notes were extracted from 100 ms pre-onset to 700 ms post-onset of the target note. Notes from melodies containing the targets (notes played in the different timbre) were not included in the analysis. All epochs were baselined to the 100 ms pre-stimulus onset period. The data was cleaned of artefacts by running wavelet enhanced independent component analysis on all of the trials from each participant separately (Castellanos & Makarov, 2006). Those

[1] The first two notes were excluded to maintain a comparable position across IC bins. As the first note of each melody had a fixed value of 5.209 and the second had a similarly high IC, the inclusion of these notes skewed the positional distribution of the IC bins such that the mean position of the bin whose IC range included these high values was much lower than for all the other bins.

[2] In the western tonal system the stability of a pitch within a key is related to its position in the hierarchy, and higher ranking/more stable pitches are often more expected than lower ranking ones (Krumhansl, 1990). In line with this and with previous reports that large pitch intervals between notes are less frequent than small ones in western melodies (Huron, 2006), increasingly high IC notes were also increasingly more tonally unstable and tended to follow larger intervals than low IC notes.

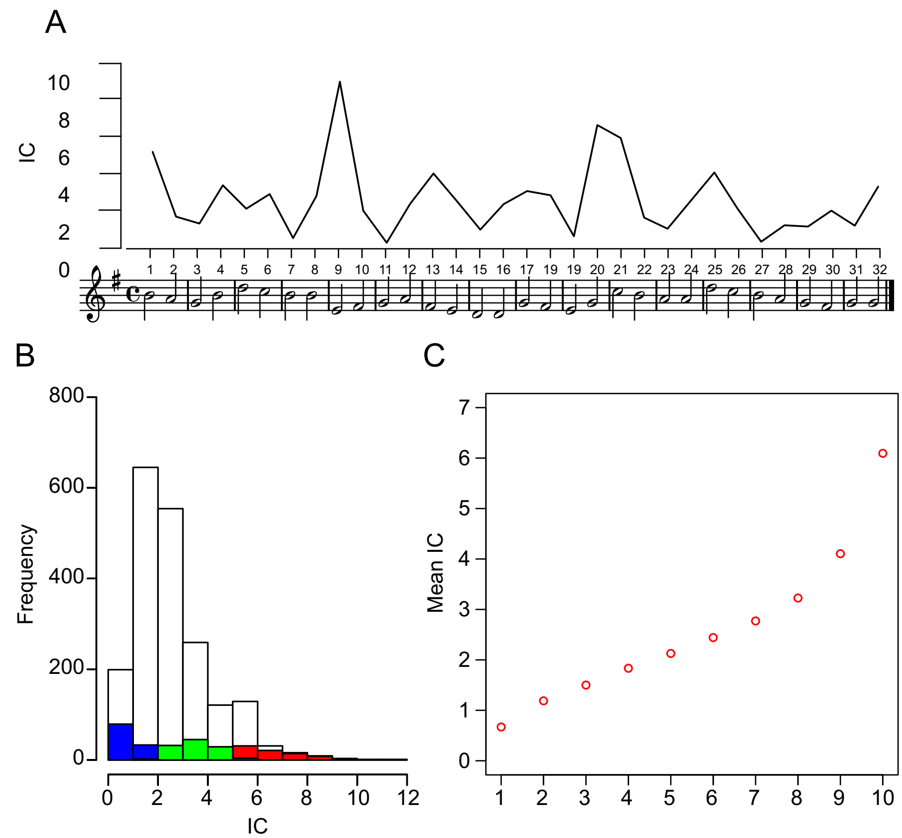

Fig. 1. (A) Information content profile of a sample melody used in the experiment. (B) The distribution of information contents for all notes in the 56 hymns (clear bars) with the distribution of probe notes in the primary analysis rendered in blue, green and red bars (low mid and high IC notes respectively). (C) The mean IC for each of the ten categories of probe notes in the secondary analysis.

Table 3
Descriptive statistics of probe notes used in the primary and secondary analyses.

			Information content	Size of preceding interval	Tonal stability	Pitch	Position
Primary analysis	Low IC	Mean	.83	1.51	4.97	68.9	17.5
		SD	.35	0.71	0.98	3.47	8.41
	Mid IC	Mean	3.40	3.29	4.53	69.35	17.86
		SD	0.87	2.04	1.29	3.41	7.89
	High IC	Mean	5.92	5.44	4.06	68.79	17.89
		SD	1.7	2.78	1.37	3.97	10.02
Secondary analysis	Bin 1	Mean	0.67	1.57	4.77	68.66	21.42
		SD	0.22	0.74	0.98	3.51	11.99
	Bin 2	Mean	1.19	1.43	4.90	67.88	20.48
		SD	0.09	0.67	1.20	3.24	11.10
	Bin 3	Mean	1.50	1.39	4.79	68.47	19.68
		SD	0.10	0.89	1.10	3.27	11.98
	Bin 4	Mean	1.83	1.08	4.72	68.19	21.08
		SD	0.09	1.15	1.09	3.26	12.56
	Bin 5	Mean	2.13	1.20	4.54	68.37	20.9
		SD	0.08	1.26	1.09	3.56	11.52
	Bin 6	Mean	2.44	1.37	4.45	68.37	19.20
		SD	0.10	1.35	1.18	3.26	11.43
	Bin 7	Mean	2.77	1.50	4.42	69.19	19.51
		SD	0.10	1.47	1.17	3.33	11.70
	Bin 8	Mean	3.23	2.44	4.28	69.56	18.72
		SD	0.17	1.58	1.07	3.13	11.43
	Bin 9	Mean	4.10	3.59	4.27	68.85	19.49
		SD	0.36	1.98	1.21	3.67	9.81
	Bin 10	Mean	6.09	5.42	4.11	69.17	19.05
		SD	1.21	2.62	1.33	4.02	10.66

components that were clearly artefacts of vertical and horizontal eye movements as well as participants' heartbeats were identified and manually removed. Epochs were then sorted by probe note and averaged to obtain mean evoked responses for each type of probe note (low, medium and high IC probe notes for the primary analysis and probe notes in IC bins 1–10 for the secondary analysis).

2.5. Data analysis

In the primary analysis, which examined which components showed significant differences according to probe category, ERPs time-locked to the onset of the target note from the individual waveforms were analysed at 16 electrodes over four regions of interest: Left anterior (F1, F3, FC1, FC3), right anterior (F2, F4, FC2, FC4), left posterior (P1, P3, PO5, PO3) and right posterior (P2, P4, PO6, PO4) sites. Peak latencies for the P1, N1 and P2 components were computed, for each participant separately, as the time point of the maximum amplitude in the 0 to 100 ms time window (P1), the time point of the minimum amplitude in the 50 ms to 150 ms time window (N1) and the time point of the maximum amplitude in the 100 to 300 ms time window (P2), respectively, relative to the 100 ms baseline activity before the note onset, so that subsequent ANOVAs could be used to examine whether individuals from the two groups showed systematic differences in these latencies. Peak amplitudes were computed as the mean amplitude of a time window running from 20 ms before to 20 ms after the mean peak latency. Latencies and peak amplitudes were submitted to individual 4 way mixed ANOVAs with group (amusic, controls) as a between subject factor and probe-type (high, medium and low IC), frontality (frontal, posterior), and laterality (left, right) as repeated measures for each component separately. In the secondary analysis, observed categorical effects were examined in greater detail, focusing on the ROIs in which effects were found in the primary analysis, to verify the presence or absence of a linear relationship with IC.

3. Results

3.1. Primary analysis: Identifying correlates of musical expectation

All participants identified at least four out of the six deviants presented over the three blocks and the two groups did not differ in detection accuracy (Controls: Mean= 5.07, SD=0.85; Amusic: Mean=4.79, SD=0.83, t=0.84, p=0.41)

Fig. 2 shows the grand average waveforms for the ERP responses of amusics and controls respectively to low (blue), medium (green) and high (red) IC notes, for each of the 16 electrodes used in the statistical analyses while Fig. 3 shows the same collapsed across the four electrodes within each of the four regions of interest (left anterior, right anterior, left posterior, right posterior). Six initial 4 way ANOVAS (group × probe type × frontality × laterality) were run**:** three examining the latency of the P1, N1 and P2 and three examining the amplitude of the same components. Follow-up ANOVAs were run, where necessary, to examine any observed interactions. Fig. 4 shows bar graphs plotting the data used in the statistical analysis.

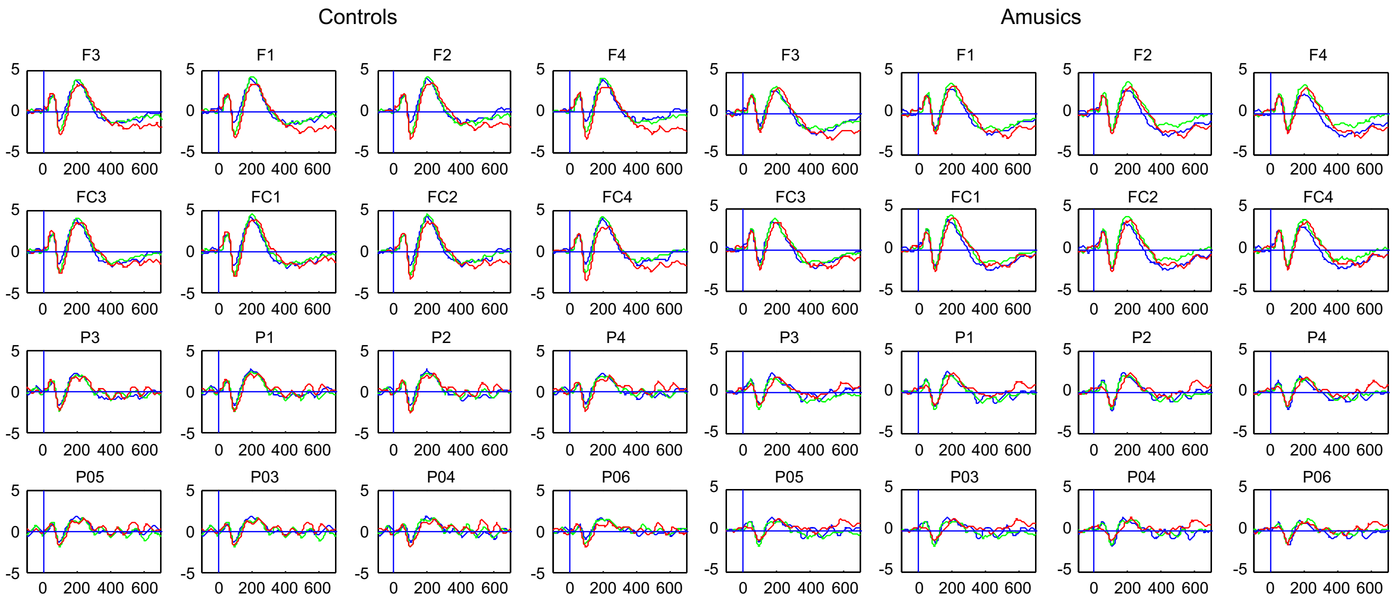

Fig. 2. Grand average waveforms for amusic and control participants for low (blue), medium (green) and high (red) IC notes, for the 16 electrodes used in the statistical analysis.

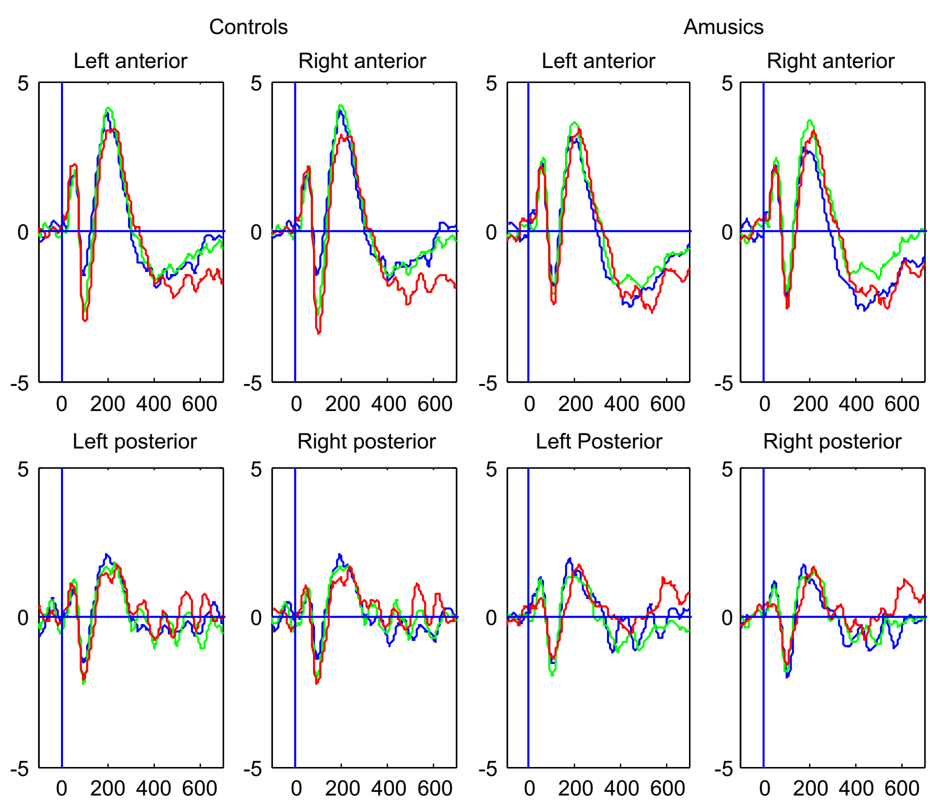

Fig. 3. Grand average waveforms for controls and amusics showing waveforms for low (blue), medium (green) and high (red) IC notes collapsed across four different regions of interest.

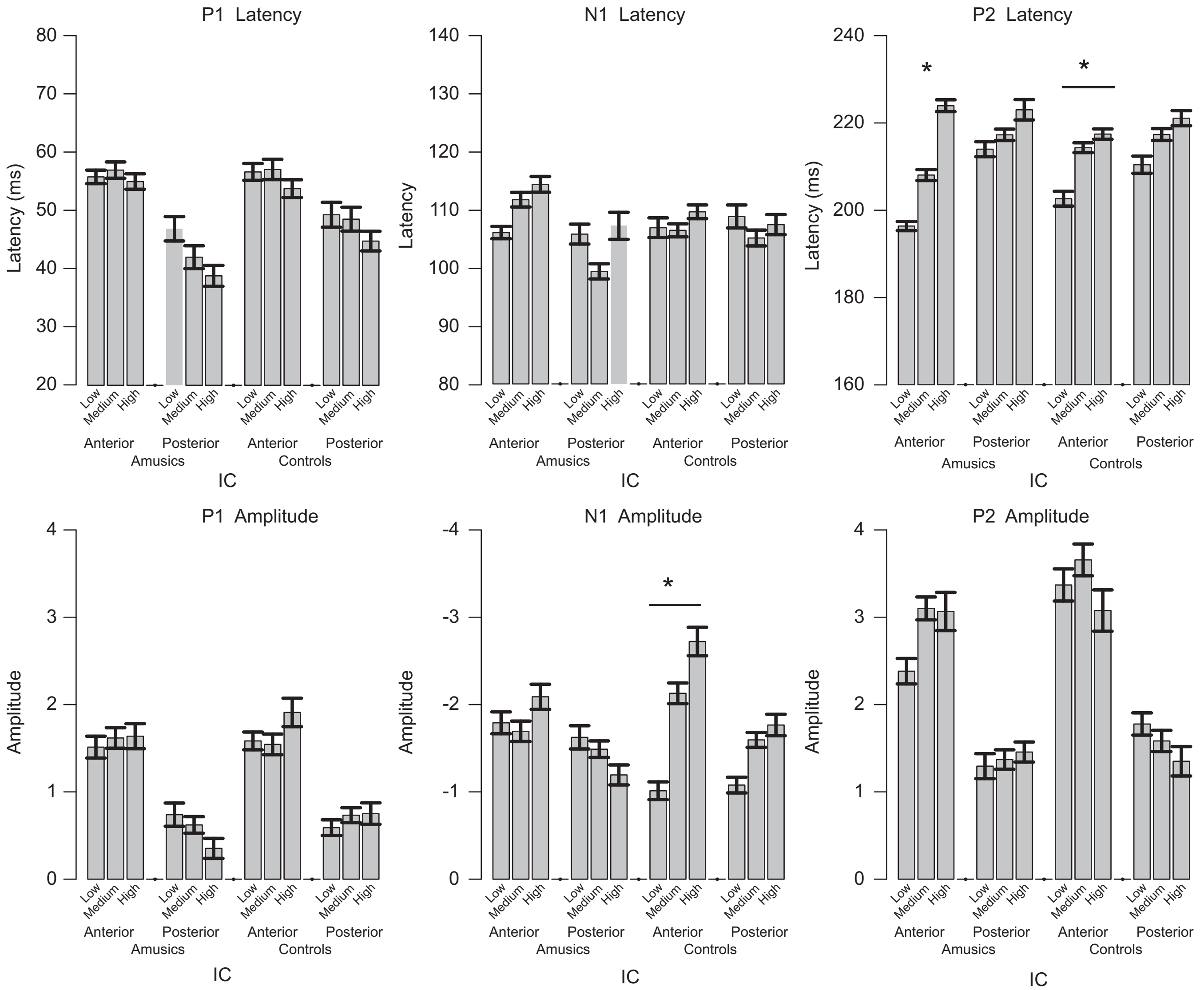

Fig. 4. Bargraphs showing P1, N1 and P2 amplitude and latency as a function of group (controls, amusic), frontality (anterior, posterior), and probe type (low, medium and high IC). Error bars indicate 1 SEM.

3.1.1. Latency

Intital 4 way ANOVAs (group × probe type × frontality × laterality) examining latency for the P1 and N1 components separately did not indicate any main effects of group or probe type or any interactions between these factors. However, for the P2 component, the 4 way ANOVA (group × probe type × frontality × laterality) revealed a significant effect of probe type, $F(2,56) = 5.52$, $p = .007$, a significant effect of frontality, $F(1,28) = 4.4$, $p = .05$, and a marginally significant interaction between probe type and frontality, $F(2,196) = 2.81$, $p = .06$. The significant main effects reflected the finding that the P2 latency for high IC events was delayed relative to that for low IC ones (low IC = 205.86 ms, mid IC = 214.16 ms, high IC = 221.4 ms) and that the P2 latency was shorter in the anterior relative to the posterior electrodes (anterior = 210.47 ms, posterior = 217.97 ms).

Follow-up 3 way ANOVAs (group × probe-type × laterality) were run on the anterior and posterior electrodes separately to explore the marginally significant interaction between probe type and frontality observed in the 4 way ANOVA. This revealed that the effect of probe type was present in the anterior ($F(2,56) = 10.65$, $p < .001$) but not the posterior electrodes ($F(2,56) = 1.10$, $p = .34$). Apart from the effect of probe type in anterior electrodes, no other main effects (laterality, group type) and none of the possible interactions (group × probe-type × laterality, probe-type × laterality, group × probe-type, group × laterality) reached significance (all $p > .1$).

3.1.2. Amplitude

Analysis of amplitudes for the P1 and P2 components did not indicate any main effects of group or probe type or any interactions between these factors. However, for the N1 component, there were significant main effects of probe type, $F(2,56) = 3.28$, $p = .045$), and frontality, $F(1,28) = 4.03$, $p = .05$, and significant interactions between group and probe type, $F(2,56) = 4.32$, $p = .018$, and between frontality and probe type, $F(2,196) = 15.8$, $p < .001$. The significant main effects of probe type and frontality reflected larger N1 amplitudes for high relative to low IC notes (low IC = −1.38 mV, mid IC = −1.73 mV, high IC = −1.94 mV) and larger N1 amplitudes in anterior than posterior electrodes (frontal = −1.91 mV, posterior = −1.46 mV) respectively, in line with the scalp map distribution seen in Fig. 5.

To investigate the significant interaction between group and probe type, follow-up one way ANOVAs of probe type were carried

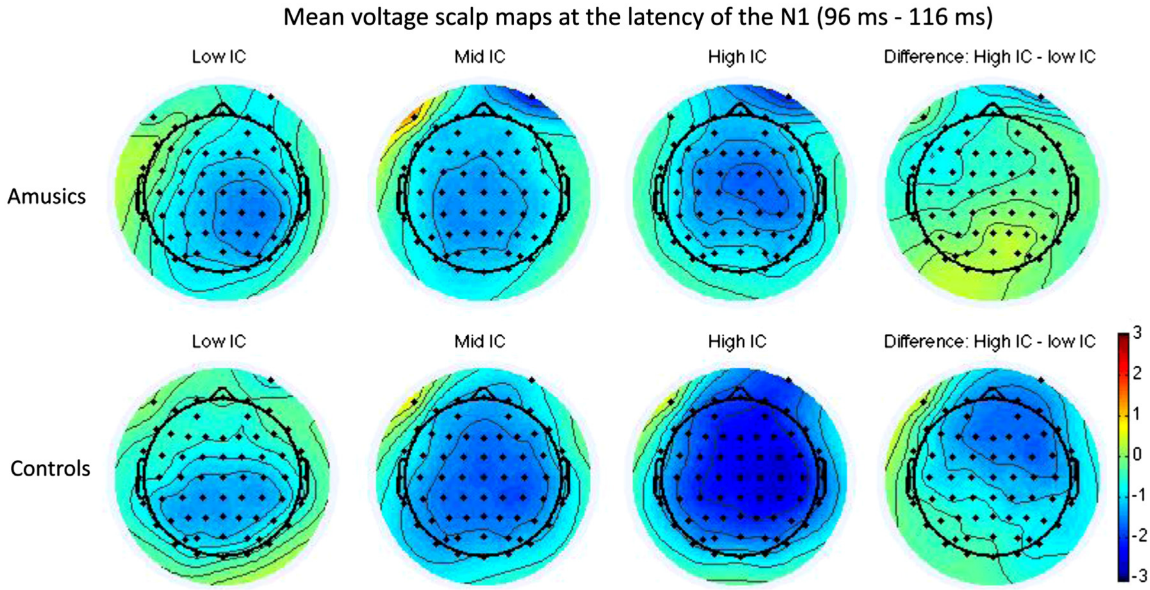

Fig. 5. Scalp maps for amusics (top row) and controls (bottom row) illustrating the early negativity effect in the N1 time window for low, medium and high IC notes, and the difference scalp map between the low and high IC conditions.

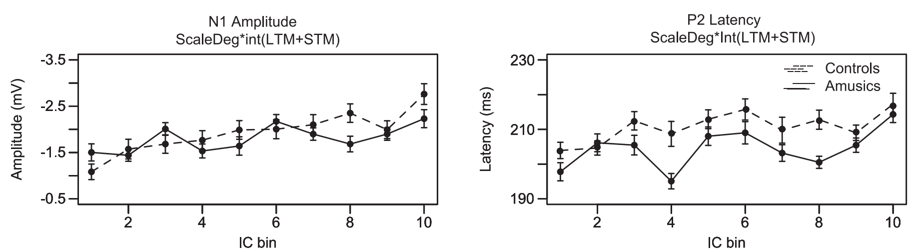

Fig. 6. Plot showing N1 amplitude and P2 latency, of the grand-averaged responses evoked over the two frontal regions of interest (left anterior: F1, F3, FC1, FC3 and right anterior: F2, F4, FC2, FC4), as a function of IC bin in control and amusic participants for the combined scale degree and interval model (LTM+STM). Error bars indicate 1 SEM.

out for each group separately. These indicated a non significant effect of probe type in both controls, $F(2,42)=1.92$, $p=.16$ and amusics $F(2,42)=.02$, $p=.98$. A follow up ANOVA, investigating the interaction between frontality and probe type, in control individuals alone, revealed a significant effect of probe type, $F(2,84)=3.16$, $p=0.04$, however, a similar ANOVA in amusic individuals alone failed to show a significant effect of probe type, $F(2,84)=0.04$, $p=0.97$, frontality, $F(1,84)=1.39$, $p=0.24$, or an interaction between the two, $F(2,84)=0.44$, $p=0.65$. Fig. 5, which plots the mean voltage scalp maps for the responses to low, medium and high IC notes and the difference in response to high IC relative to low IC notes, demonstrates the attenuated N1 amplitude effect in amusics relative to controls and shows the localization of the effect to anterior rather than posterior areas.

To summarize, in the primary analysis, two main effects were seen in controls in response to unexpected notes namely a longer latency P2 and a larger N1 at frontal scalp locations. Amusic participants showed the former but not the latter effect.

3.2. Secondary analyses: Examining and characterizing the relationship between observed effects and IC

Further analysis sought to examine the strength and nature of the frontally maximal early negative response (increase in N1 amplitude with increasing IC) and the P2 latency effect (increase in P2 latency with increasing IC) observed in the primary analyses. Using more data points per category and also a greater number of categories (see Section 2.2.2), the secondary analysis also provided a more sensitive treatment of the data relative to the primary analysis. Fig. 6 illustrates how the N1 amplitude and P2 latency varied as a function of IC bin for the IC profile used in the primary analyses (ScaleDeg*Int: LTM+STM). Significant correlations were found in controls for N1 amplitude ($r(8)=-0.9$, $p<.001$) and for P2 latency ($r(8)=0.66$, $p=.04$) providing further support for the earlier observed relationships in the primary analyses and evidence that these relationships are linear. In amusics, a relationship was also observed between P2 latency and IC in line with the primary analyses ($r(8)=0.59$, $p=0.07$). However, in contrast to the primary analysis, which revealed no significant effect of probe type, this secondary analysis revealed a significant relationship between N1 amplitude and IC in amusic individuals ($r(8)=-0.66$, $p=.04$), although the lower significance for this group relative to controls suggests that the extent to which the amusics were processing IC was reduced in extent relative to controls.

Compared to the relatively smooth trajectory in N1 amplitude as IC increased in controls, the relationship between these two variables showed a degree of unevenness in amusics. This rather unsystematic relationship between N1 amplitude and IC in amusics may partly explain the lack of significance for this group in the primary analysis. Thus while the primary analysis suggested a complete lack of sensitivity of the N1 amplitude to IC, the secondary analysis suggested a more nuanced view, whereby amusics appear to be sensitive to IC in music but to a lesser extent than controls.

3.3. Relationship with behaviour

Analysis was carried out to investigate whether the amplitude of the N1 and latency of the P2 components could be predicted by the participants' performance on the implicit and explicit tasks, which we reported in Omigie et al. (2012). The implicit and explicit tasks from this previous study were carried out in a separate testing session approximately 6 months before the current experiment. The stimuli used in the current experiment were identical to the stimuli used in the implicit task of that study and 10 controls and 8 amusics took part in both the current and previous study. For each of these participants, the t statistic of t-tests contrasting the amplitude of the N1 peak for low and high IC notes from the primary analysis, and those contrasting the latency of the P2 peak for low and high IC notes (also from the primary analysis) served as measures of the strength of these effects in each of these individuals. Similarly, the t statistic from t-tests contrasting the ratings given to the low and high IC notes in the explicit task served as a measure of the strength of explicit knowledge, while the t statistic contrasting reaction times to low and high IC notes in the implicit task served as a measure of the strength of implicit knowledge.

The results of correlational analyses can be seen in Table 4. Taking both groups together, N1 amplitude was shown to correlate with performance on the explicit task ($r(16)=-0.54$, $p=0.02$). This indicated that the greater the N1 amplitude was for high IC relative to low IC events, the more unexpected listeners found high IC events relative to low IC ones. The P2 latency showed a similar trend ($r(16)=-0.73$, $p=0.08$). Put simply, the greater the P2 latency was for high IC relative to low IC events, the greater the participants' unexpectedness ratings were for high IC relative to low IC ones.

On a group level, the controls showed a tendency towards the same result with regard to N1 amplitude ($r(8)=-0.56$, $p=.09$) but this was not the case for the amusics ($r(6)=-0.18$, $p=.67$). Finally, neither the N1 amplitude effect nor the P2 latency effect correlated with performance on the implicit task either at a group or a combined level.

Further analysis explored the extent to which the current EEG measures correlated with performance on the MBEA pitch subtests and psychophysically measured pitch thresholds. As before, the correlation between the EEG measures (N1 amplitude and P2 latency effects) shown at an individual level and the scores on the MBEA subtests and pitch thresholds are given in Table 5. When looking at both groups together, the N1 amplitude effect correlated with performance on the interval and contour subtests ($r(16)=0.52$, $p=0.03$ and $r(16)=.47$, $p=0.04$, respectively) and marginally with the scale subtest ($r(16)=0.42$, $p=.08$), once again showing a relationship between the N1 amplitude effect and musical competence. However, while the P2 latency effect did not show a significant relationship with the MBEA subtests, or pitch detection thresholds in either amusics or controls, it is interesting to note that it correlated with pitch direction thresholds in amusics ($r(16)=-0.77$, $p=0.03$). This relationship was absent in controls although this might be due to the reduced variance in the pitch direction thresholds of controls compared to amusics (amusics: variance=2.65, controls variance= 0.01).

Table 4
Results of Pearson correlations of the N1 amplitude and P2 latency effects with performance on the implicit and explicit judgment tasks (reported in Omigie et al., 2012).

			Implicit task	Explicit task
N1 amplitude	Amusics	$r(6)$	0.27	−0.18
		p	0.52	0.673
	Controls	$r(8)$	−0.40	−0.56
		p	0.26	0.09
	Both	$r(16)$	−0.12	**−0.54**
		p	0.62	**0.02**
P2 latency	Amusics	$r(6)$	−0.42	−0.24
		p	0.29	0.56
	Controls	$r(8)$	−0.40	−0.10
		p	0.26	0.78
	Both	$r(16)$	0.29	−0.73
		p	0.24	0.08

3.4. Timecourse of processing

A further set of analyses asked whether listeners showed evidence of responding differently to IC values as the music unfolded. To explore this, melodies were split into two halves and 20% of notes in each half were assigned to a 'low IC' and 20% to a 'high IC' category. A three way ANOVA (probe type (low, high IC) × time (early, late) × group (controls, amusics)) of N1 amplitude revealed a main effect of Probe type ($F(1,28)=8.92$, $p<.01$) confirming previous findings, and a main effect of time ($F(1,28)=6.75$, $p=0.05$) indicating that the amplitudes of the N1 evoked in the second half of the melody were generally smaller. A 3 way ANOVA (probe type × time × group) of P2 latency, indicated a marginally significant effect of Probe type ($p=0.1$). However, there was no interaction of time with probe type in either the ANOVA exploring N1 amplitude or that exploring P2 latency. This indicates that the relationship between IC and the participants' electrophysiological responses did not change over the course of hearing the melody and that the model's predictions accounted adequately for any potential effects of time. Fig. 7 shows the mean IC and

Table 5
Results of Pearson correlations of the N1 amplitude and P2 latency effects with performance on the pitch subtests of the MBEA and psychophysically measured pitch detection and discrimination thresholds.

			MBEA scale	MBEA interval	MBEA contour	Pitch Detection Ćthresholds	Pitch Discrimination ĆThresholds
N1 amplitude	Amusics	$r(6)$	−0.06	0.15	0.19	0.03	−0.33
		p	0.88	0.72	0.65	0.95	0.43
	Controls	$r(8)$	0.21	0.37	0.25	−0.45	−0.17
		p	0.55	0.30	0.49	0.19	0.63
	Both	$r(16)$	0.42	**0.52**	**0.47**	−0.26	−0.35
		p	0.08	**0.03**	**0.04**	0.29	0.14
P2 latency	Amusics	$r(6)$	0.24	−0.31	−0.24	−0.57	**−0.77**
		p	0.56	0.45	0.56	0.14	**0.03**
	Controls	$r(8)$	0.04	−0.04	−0.05	0.09	−0.44
		p	0.91	0.9	0.87	0.80	0.20
	Both	$r(16)$	0.17	−0.21	−0.21	−0.03	0.04
		p	0.50	0.39	0.40	0.9	0.89

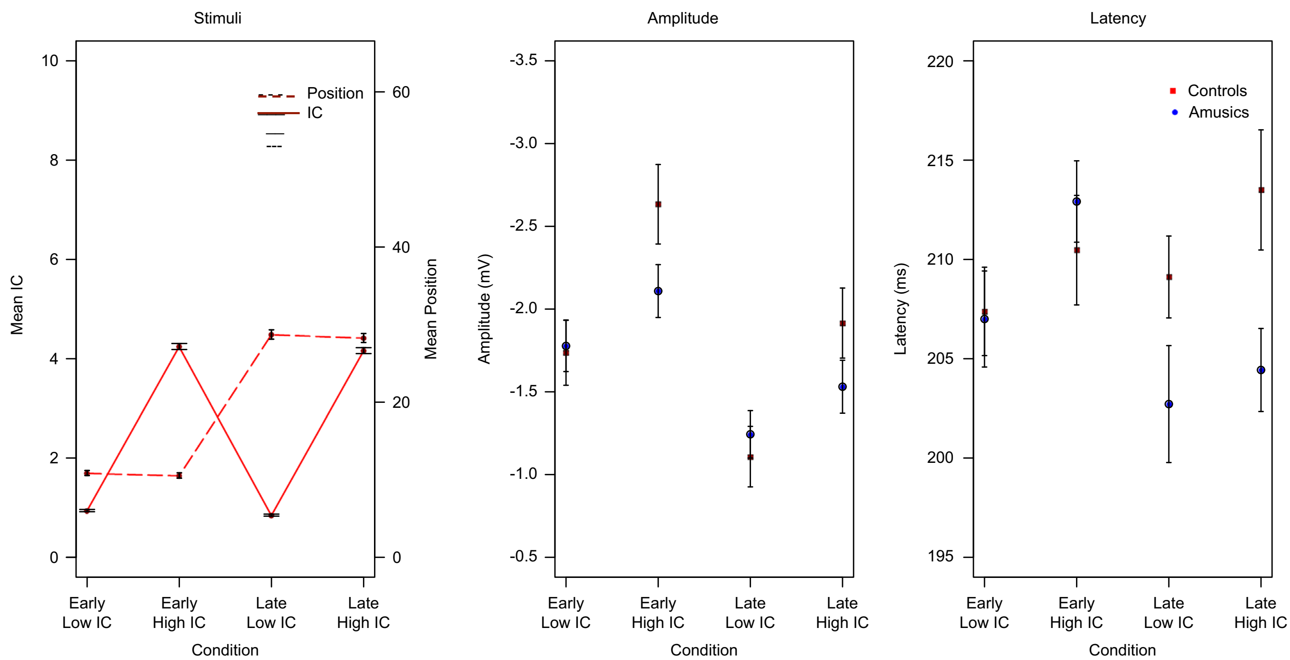

Fig. 7. Time course of processing stimuli and results. (A) Plot showing mean IC and position of the probe notes in the four categories (early low IC, early high IC, late low IC and late high IC. (B) Mean N1 amplitude for the two groups across the four note categories. (C) Mean P2 latency for the two groups across the four note categories.

positional distribution of the four categories as well as the mean N1 amplitude and P2 latency observed for each of them.

3.5. *Supplementary analyses*

3.5.1. *Examining the effects of per-melody IC binning*

The mode of binning reported in the secondary analysis (3.2) is based on notion that a listener's response to a note is best predicted by the absolute IC of the note regardless of the mean IC of the melody in which it is found. Here we carried out a different type of binning to explore the evidence for an alternative scenario, whereby listeners treat each melody as individual objects and adapt physiologically to them. Melodies vary in the mean level of IC of the constituent notes (range$=$1.65 to 3.43), and one possibility is that listeners adjust to the level of surprise in each melody. Thus, in this second method of binning, each melody was considered a set (resulting in 56 sets) and notes in each set were sorted in order of increasing IC before being allocated to 10 different bins based on their relative position within the set (each melody).

The main difference between the binning methods is clearly shown in Supplementary Fig. 1A (note that error bars reflect 40 SEM for the sake of visibility) where it can be seen that while the mean IC per bin was very similar across binning methods, binning per melody resulted in bins with highly overlapping IC values, compared to when binning across melodies, where the range of ICs in the 10 bins did not overlap at all. In the first scenario, where a listener's response to a note is best predicted by the absolute IC of the note, one would expect a stronger relationship between neural signatures of expectation formation and the mean IC values of bins collapsed across melodies. However, in the alternative scenario, where listeners may adapt physiologically to the mean IC of a melody, one would expect a stronger relationship between neural signatures of expectation formation and the mean IC values of bins arrived at by binning within melodies.

Results from this second type of binning are presented in Supplementary Fig. 3A–C. A first interesting finding was that the relationship was stronger for both groups when notes were binned per melody than when binned across melodies (See Supplementary Table 1: Controls: 11 significant correlations for per-melody binning vs 10 for across-melody binning; Amusics: 10 significant correlations for per melody binning versus 7 for across melody binning). This lends weight to the proposal that individuals adjust to the mean IC of a melody. However, it is important to note that this effect is a subtle one since the significant relationships ($p<.05$) in the per melody binning did not have a correlation coefficient that was significantly different from the nonsignificant ones in the across melody relationship (To test whether two correlation coefficients differed in strength the procedure developed by Fisher (1921) was used).

A second interesting finding was that the difference in the two methods of binning was seen especially in the amusic participants, and particularly in the P2 latency effect. Specifically, while the majority of correlations between the P2 latency effect and IC output were nonsignificant when the melodies were binned across melody, almost all the correlations were shown to be significant or approach significance when notes were binned per melody. This would seem to suggest that the novel P2 latency effect found here reflects a response to unpredictable musical events that is scaled to the overall unpredictability of the musical context. However, here again, while the per melody correlations tended to reach significance more often than the across melody correlations, the corresponding correlation coefficients in the two type of analysis (across and per melody binning) did not differ in strength (all $p>0.1$). Unfortunately, the current design does not allow any further testing of the hypothesis that the novel P2 latency effect found here is scaled to the overall unpredictability of the musical context, but we feel further experimentation in controls to clarify the P2 latency mechanism and how it is modulated, will be useful in explaining the differences between the groups seen here.

3.5.2. *Examining the effects of different model configurations*

While previous work on the typical listening populations (Pearce et al., 2010) led to our choice of model version (ScaleDeg-*Int, combined LTM+STM) in the main analyses, it was also of considerable interest to ask which components of the model may be driving the observed ERP effects and whether these were similar or different for amusics and controls. In particular, the STM reflects short-term statistical learning from the current musical context while the LTM reflects schematic effects on expectations acquired through long-term listening to music. Orthogonally, it is possible to vary the representations (viewpoints) of the musical surface, allowing us to analyse the output of the model for Pitch Interval and Scale Degree separately.

Thus, IC profiles generated by different configurations of the model, based on the Scale Degree and Interval features, STM and LTM, were also correlated with the ERP effects observed here. Supplementary Fig. 1B shows the positional distribution of notes in the melodies that were assigned to the different bins in these analyses and demonstrates the similarity of position across these bins. The results of the correlations between mean IC and the N1 amplitude and P2 latency effects for each of these analyses are detailed in Supplementary Table 1, and are plotted in Supplementary Figs. 2A–C and 3A C.

While these analyses revealed a complex picture that would benefit from further experimentation, several points are worth noting. Perhaps one of the most important of these is that the N1 effect in controls is more sensitive to scale degree than the N1 response in amusics. Specifically a comparison of corresponding correlation coefficients across the two groups revealed that the only significant differences were in the correlations between (i) the N1 effect and scale degree (STM+LTM), (ii) the N1 effect and scale degree (LTM only), and (iii) the N1 effect and Interval(LTM only), whereby these relationships were significantly greater in controls than in amusics. On a related note, however, it is important to note that while these results would seem to suggest that amusics are largely insensitive to scale degree, there was nevertheless a good relationship between scale degree (LTM and STM) and the P2 latency effect. This indicates that amusic individuals are also sensitive to tonal features of the musical surface even if to a lesser degree: in one response (the P2) rather than two (the N1 and P2) as in controls.

Another point worthy of note is that, in controls, the P2 latency effect did not show the same degree of sensitivity to IC as the N1 amplitude effect, in terms of the number of versions of the model eliciting the effect. Specifically, while the vast majority of model versions yielded the N1 effect, only the model versions in which the Scale degree and Interval features were combined yielded the P2 latency effect. This suggests that this effect reflects expectations based specifically both on melodic (Intervallic) and tonal (Scale Degree) structure. Finally, we propose that the fact that the LTM has a stronger effect than the STM, especially for Scale Degree models, may be taken as evidence that long-term schematic effects of tonal listening have more robust effects on expectation than short-term effects, at least for this corpus.

4. Discussion

One of the defining characteristics of individuals with congenital amusia is difficulty in the detection of gross musical violations. In a previous study, two versions of a melodic priming paradigm and the predictions of the current computational model of melodic expectation were used to examine how amusic listeners responded to notes of high or low IC in the context of ecologically valid melodies (Omigie et al., 2012). This study indicated that amusic individuals and controls were similar in the extent to which they showed evidence of implicit musical expectations but critically, that amusics were significantly worse than controls at using explicit ratings to differentiate between high and low IC events in a melodic context.

The current study used EEG recordings to further investigate the observed discrepancies between the implicit and explicit music anticipatory capacities in those with amusia. Analysis revealed an unanticipated effect of IC on P2 latency, whereby high IC notes reliably elicited a delayed P2 component. This effect appeared to a similar extent in amusics and controls. However, a second effect of IC, which we anticipated based on previous literature on the ERP correlates of musical expectation, dissociated the two groups. Notes with high IC were characterized by an early frontal negativity (similar to that often referred to as the ERAN). Importantly, this effect was found to be much less robust in amusics, to the point of failing to emerge in the context of more traditional ERP analyses employing a limited number of event categories (as opposed to one employing many as in the secondary analysis).

Critically, the predicted finding of a diminished early frontal negativity in amusic individuals is in line with a previous ERP study which investigated melodic processing in amusia (Peretz et al., 2009). This study showed the absence of an N200 in response to out-of-key notes, (to which the high IC notes investigated here may be compared). However, as in the present study, a degree of sensitivity to such notes was, nevertheless, also demonstrated since the authors reported that the N200 was marginally significant at the F4 electrode, in contrast to controls in whom significance was reached in several electrodes. The finding, here, of a significant correlation between the size of the early negative response and IC in our amusic participants similarly demonstrates that amusics are sensitive to less probable events in a musical context, albeit to a reduced extent compared to controls.

This diminished early frontal negativity in the neural responses of amusic individuals is also congruent with the notion that early mechanisms may show a relationship with the degree of musical expertise a listener possesses. Koelsch et al. (1999) conducted a study where expert violinists and musical novices were presented with an oddball sequence in which perfect major chords (standard stimuli) were interspersed with the same chords with a slightly mistuned centre tone (the deviant stimuli). In a passive condition, in which participants engaged in reading a self selected book and were told to ignore the stimuli, these authors reported that the musicians showed a large MMN to deviant chords while in contrast the novices did not. In an active detection condition, in which musicians detected 83% of the stimuli and the novices only 13%, the authors reported that the musicians showed the equivalent of a very strong MMN like response, while the novices showed a response that was slightly larger than in the passive condition but still much smaller than in the musicians. The authors argued that the superior ability of the musicians to consciously detect the slightly impure chords was reflected in the much larger response they showed relative to novices who were less able to detect these deviants. In another study (in which listeners listened attentively for timbral deviants rather than musical deviants as in the current study), Koelsch et al. reported that musical experts showed a larger ERAN than novices to harmonically inappropriate chords in the context of a chord progression (Koelsch et al., 2002). Once again they speculated that this might be because musicians have more specific expectations of how music should unfold due to greater explicit knowledge of the theory of musical harmony (Bharucha, 1984). In a follow-up study (Koelsch et al., 2007), they were able to provide support for this theory of a relationship between explicit knowledge and the ERAN amplitude by showing that in addition to producing a larger ERAN, musicians were indeed more accurate than non-musicians at identifying irregular endings to a chord progression

Extrapolating from the findings of Koelsch et al., we suggest that the current finding of a diminished early negativity in amusic relative to controls is evidence that they have internalized the regularities in music but have a less robust representation of this information which will often fail to reach conscious awareness. This interpretation is supported by results from a behavioural study showing that individuals with amusia are just as capable as controls of internalising transition probabilities in novel tonal materials even though they show much less confidence in their decisions as well as inferior explicit knowledge of how they perform (Omigie & Stewart, 2011).

It is also interesting to relate the current findings of a diminished but present sensitivity to IC, as indexed by the N1, to the diminished but present explicit knowledge of musical structure found in amusics in the explicit task reported in Omigie et al. (2012). The relationship between the two is evident in our finding of a correlation between N1 amplitude and performance on the explicit task. Indeed taken together, the study seems to suggest that in terms of conscious access to musical knowledge, the difference between congenitally amusic and typical individuals, may not be a purely categorical one. Indeed it suggests that rather than being an 'all or none' phenomenon, the ability to report on the structure of musical events may be graded. This is in line with theories that suggest that implicit and explicit knowledge are not separate phenomena but rather that intact performance on implicit tasks such as the reaction time task used in the behavioural study of 2012 (Omigie et al., 2012), indicates the presence of some, if not complete, levels of knowledge (Cleeremans & Jimenez, 2002). Following in this vein, the current data may be interpreted as suggesting that amusic individuals are not categorically different from controls in terms of their levels of musical awareness (in an all or none sense), but lie lower on the spectrum of possible degrees of awareness.

In terms of the timing of the observed effects, while a response at the latency of the N1 may seem rather early to be associated with explicit knowledge, it has been argued that early mechanisms do play an important role in the emergence of conscious evaluation of less probable events in the auditory environment (Naatanen, 1990) with the theory holding that these early mechanisms possess attention-triggering properties (Naatanen, 1990; Winkler, 2007). This theory, along with previous work from Koelsch et al. (1999, 2002, 2007), supports our suggestion that the attenuated early negative response seen in amusic individuals may be related to, or even underpin, their reduced ability to explicitly detect notes with high IC. This interpretation is also in line with a paper exploring melodic processing in individuals with a disorder known as tune deafness (related to amusia but assessed using a different diagnostic battery) (Braun et al., 2008). The authors observed that wrong notes inserted in familiar melodies elicited a mismatch negativity in controls and not in tune deaf subjects, but nevertheless evoked a robust P300 response in both tune deaf individuals and control participants. In this study, participants were asked to the listen to the melodies but not specifically asked to detect abnormal notes.

These results from Braun et al. present an interesting parallel with our somewhat paradoxical finding that the neural response mostly likely to reflect implicit knowledge (P2 latency) occurs later than the component (N1 amplitude) that is likely to index explicit knowledge. Critically, these authors also interpreted the lack of the MMN in tune deaf individuals as a marker of the lack of awareness of musical deviants, and suggested that computations in the early auditory areas play a role in determining whether deviant auditory information will be consciously perceived. Like those authors, we suggest that the preservation of one effect in the absence of the other may arise from the fact that the sources of the evoked responses are different. Braun and colleagues suggest that tune deafness may be comparable to blindsight whereby information bypasses the early mechanisms that regulate conscious perception but is processed in later areas. In the current study, we show that this finding may generalize to amusia. Further, we show that an attenuated early negativity may occur not just in response to veridical melodic deviants or artificial inserted schematic violations but also to high IC notes in the context of natural melodies without alteration.

An interesting additional finding was that of a significant influence of note IC on the latency of the P2. While numerous studies have examined the neural correlates of musical expectation (Besson & Macar, 1987; Besson & Faïta, 1995; Paller, McCarthy, & Wood, 1992; Verleger, 1990), to our knowledge, the current study is the first report of a P2 latency effect. It has been suggested that the latency of certain ERP components (N100, P200, N200, P300) is an indication of the speed with which stimuli are evaluated (Polich, Ellerson, & Cohen, 1996) and indeed, the latency of several ERP components has been shown to co-vary with task difficulty, whereby more complex tasks result in longer latencies of the P1, N1, P2 and P3 (Goodin, Squires, & Starr, 1983). In the current study, participants were required to evaluate each note for a change in timbre and we interpret the P2 latency effect observed here to reflect the greater difficulty participants had in processing the timbre of unexpected notes relative to expected ones. Another possibility that may be proposed is that the delayed P2 is a result of slower recovery from a deeper N1, however the fact that only the N1 component showed a group effect speaks against this interpretation. Indeed if the observed P2 latency were simply a side effect of a deeper N1, one would not expect to see the dissociation between the two that is observed when the profiles are examined in isolation in amusics.

We suggest that the current finding in amusics of a delayed P2 latency for high IC relative to low IC notes, (as was found in controls) is interesting in showing a degree of sensitivity to IC in amusics. While it is tempting to describe the P2 latency effect seen here as a marker of implicit knowledge, its failure to correlate with the performance on the implicit task in our previous behavioural experiment (Omigie et al. 2012) raises the question of its precise behavioural correlate. However, it is important to note that the two paradigms differ in terms of motor and attentional task demands. Specifically, while listeners had to evaluate notes in terms of their timbral quality in both studies, the reaction time task required listeners to make responses to notes which were clearly indicated by the hand of a clock counting down to the target note while in the EEG task, listeners were not directed to the notes to be evaluated but had to attend to every note in each melody. Further while both paradigms required responses to be made, there was a speeded response component in the behavioural but not in the EEG study. At any rate, it remains important to explore this P2 effect further in a group of control individuals and under different conditions and task requirements before attempting to draw any conclusions regarding its meaning.

Finally, it is worth noting the results of the exploratory analysis which addressed the degree to which listeners show sensitivity to the mean levels of IC in a melody and also the extent to which they (amusic and control listeners) rely on different musical features when forming musical expectations. While preliminary, some specific findings are of interest in formulating hypotheses to test in future research. Firstly, the finding that there is a stronger relationship in both groups between mean IC and the ERP effects when the notes are binned per melody versus when they are binned across melodies (Supplementary Table 1/Supplementary analyses: 3.5.1) suggests that individuals are indeed sensitive to the mean level of IC in a melody—they scale their expectations according to the overall expectedness of the musical context. That this was specifically the case for the P2 latency effect seems to suggest that this novel effect reflects a response to unpredictable musical events that is scaled to the overall unpredictability of the musical context, although this requires further experimentation. Secondly, although the finding that the effects of Scale Degree on N1 amplitude are stronger for controls than for amusics (Supplementary Table 1/Supplementary analyses. 3.5.2) might be taken to suggest that amusics are not sensitive to the tonal representations of melody, the finding that the P2 latency effect in amusics shows a strong relationship with Scale degree IC suggests further work is needed to explore how amusics differ from controls in terms of cognitive and neural representations of musical knowledge.

In sum, the current electrophysiological study, provides an important extension to the published behavioural studies that have reported diminished explicit awareness of musical deviance in amusia. Specifically, we report an attenuated early negative response to unexpected notes in amusic individuals. Given the established link between the amplitude of early negative deflections and explicit knowledge of musical deviance in typical listeners (e.g. Koelsch et al., 2007; Miranda & Ullman, 2007), we suggest this finding provides a potential biological correlate of the musical perceptual deficits seen in this group.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2013.05.010.

References

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

Besson M. and Macar F., An event-related potential analysis of incongruity in music and other non-linguistic contexts, Psychophysiology 24, 1987, 14–25.

Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1278–1296.

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, *16*, 485–518.

Braun, A., McArdle, J., Jones, J., Nechaev, V., Zalewski, 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*(6), 2349.

Castellanos, N. P., & Makarov, V. A. (2006). Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis. *Journal of Neuroscience Methods*, *158*(2), 300–312.

Chaplin, J. P. (1985). *Dictionary of psychology* (2nd ed.). New York: Laurel rev..

Cuddy, L. L. & Lunny, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgements of continuity. Perception and Psychophysics, 57, 451–462.

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, computational and philosophical consensus in the making.* Hove, UK: Psychology Press.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., & Spector, T. (2001). *Genetic correlates of musical pitch recognition in humans Science*, *291*, 1969–1972.

Fisher, R.J (1921). On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, *1*, 3–32.

Goodin, D. S., Squires, K. C., & Starr, A. (1983). Variations in early and late event-related components of the auditory evoked potential with task difficulty. *Electroencephalography and clinical neurophysiology*, *55*(6), 680–686.

Henry, M., & McAuley, D. J. (2010). On the prevalence of congenital amusia. *Music Perception*, *27*, 413–418.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.

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.

Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: when less is better than more. *The Journal of Neuroscience*, *27*(47), 13028–13032.

Juslin, P. N., & Vastfjall, D. (2008). Emotional responses to music: the need to consider underlying mechanisms. *Behavioural Brain Science*, *31*, 559–575.

Kalmus, H., & Fry, D. B. (1980). On tune deafness (dysmelodia): frequency, development, genetics and musical background. *Annals of Human Genetics*, *43*(4), 369–382.

Kim, S. G., Kim, J. S., & Chung, C. K. (2011). The effect of conditional probability of chord progression on brain response: an MEG study. *PLoS ONE*, *6*(2), 9.

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, *10*(6), 1309–1313.

Koelsch, S., Gunter, T., Friederici, A, & Schröger, E. (2000). Brain indices of music processing: "nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541.

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: an event-related brain potential study. *Psychophysiology*, *39*(5), 657–663.

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: an ERP study of music perception. *Psychophysiology*, *44*(3), 476–490.

Koelsch, S., & Jentschke, S. (2010). *Differences in electric brain responses to melodies and chords Journal of Cognitive Neuroscience*. 22. MIT Press2251–2262.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Research*, *1142*, 169–177.

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain: a Journal of Neurology*, *133*, 1682–1693.

Loui, P., Grent-'t-Jong, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Brain Research. Cognitive Brain Research*, *25*(3), 678–687.

Loui, P., Alsop, D., & Schlaug, G. (2009a). Tone deafness: a new disconnection syndrome? *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(33), 10215–10220.

Loui, P., Wu, E. H., Wessel, D. L., & Knight, R. T. (2009b). A generalized mechanism for perception of pitch patterns. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *29*(2), 454–459.

Mackay, DJC. (2003). *Information theory, inference, and learning algorithms*. Cambridge, UK: Cambridge University Press.

Mandell, J., Schulze, K., & Schlaug, G. (2007). Congenital amusia: an auditory-motor feedback disorder? *Restorative neurology and neuroscience*, *25*(3-4), 323–334.

Marmel, F., Perrin, F., & Tillmann, B. (2011). Tonal expectations influence early pitch processing. *Journal of Cognitive Neuroscience*, *23*, 3095–3104.

Miranda, R., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: an event-related potential study. *NeuroImage*, *38*(2), 331–345.

Moreau, P., Jolicoeur, P., & Peretz, I. (2009). Automatic brain responses to pitch changes in congenital amusia. *Annals of the New York Academy of Sciences*, *1169*, 191–194.

Narmour. (1990). The Analysis and cognition of basic melodic structures: The implication-realization model. Chicago: University of Chicago Press.

Naatanen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive functions. *The Behavioural and Brain Sciences*, *13*, 201–288.

Naatanen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Laurence Erlbaum.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544–2590.

Nicholson, S., Knight, G. H., Dykes, & Bower, J. (1950). *Ancient and modern revised*. Suffolk, UK: William Clowes and Sons.

Omigie, D., & Stewart, L. (2011). Preserved statistical learning of tonal and linguistic material in congenital amusia. *Frontiers in Psychology*, *2*, 109.

Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia*, *50*, 1483–1493.

Paller, K. A., McCarthy, G., & Wood, C. C. (1992). Event-related potentials elicited by deviant endings to melodies. *Psychophysiology*, *29*(2), 202–206.

Pearce, M. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition*. City University, London, UK: Doctoral Dissertation, Department of Computing.

Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: the influence of context and learning. *Music Perception*, 377–405.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, *50*(1), 302–313.

Peretz, I., Ayotte, J., Zatorre, R., Mehler, J., Ahad, P., Penhune, V. & Jutras, B. (2002) Congenital Amusia: A Disorder of Fine-Grained Pitch Discrimination. *Neuron*, vol. 33, pp. 185–191.

Peretz, I., Champod, A. S., & Hyde, K. (2003). *Varieties of musical disorders. The Montreal Battery of Evaluation of amusia*. 999. Annals of the New York Academy of Sciences58–75.

Peretz, I., Brattico, E., & Tervaniemi, M. (2005). Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology*, *58*(3), 478–482.

Peretz, I., Cummings, S., & Dubé, M. P. (2007). The genetics of congenital amusia (tone deafness): a family-aggregation study. *American Journal of Human Genetics*, *81*(3), 582–588.

Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: in tune, out of key, and unaware. *Brain: A Journal of Neurology*, *132*, 1277–1286.

Polich, J., Ellerson, P. C., & Cohen, J. (1996). P300, stimulus intensity, modality, and probability. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *23*(1–2), 55–62.

Schellenberg, E. G. (1997). Simplifying the implication-realization model of melodic expectancy. *Music Perception*, *14*(3), 295–318.

Schmuckler, M. A. (1997). Expectancy effects in memory for melodies. *Canadian Journal of Experimental Psychology*, *51*, 292–306.

Schulkind, M. D., Posner, R. J., & Rubin, D. C. (2003). Musical features that facilitate melody identification: How do you know it's your song when they finally play it? *Music Perception*, *21*, 217–249.

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

Tillmann, B., Gosselin, N., Bigand, E., & Peretz, I. (2012). Priming pardigm reveals harmonic structure processing in congenital amusia. *Cortex*, *48*, 1073–1078.
Verleger, R. (1990). P3-evoking wrong notes: unexpected, awaited, or arousing? *The International Journal of Neuroscience*, *55*(2–4), 171–179.
Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia: beyond a fine-grained pitch discrimination problem. *Memory*, *18*(6), 657–669.
Williamson, V. J., Liu, F., Peryer, G., Grierson, M., & Stewart, L. (2012). Perception and action de-coupling in congenital amusia: sensitivity to task demands. *Neuropsychologia*, *50*(1), 172–180.
Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*, 147–163.