



HAL
open science

Consonance Perception in Congenital Amusia: Behavioral and Brain Responses to Harmonicity and Beating Cues

Jackson Graves, Agathe Pralus, Lesly Fornoni, Andrew Oxenham, Barbara
Tillmann, Anne Caclin

► **To cite this version:**

Jackson Graves, Agathe Pralus, Lesly Fornoni, Andrew Oxenham, Barbara Tillmann, et al.. Consonance Perception in Congenital Amusia: Behavioral and Brain Responses to Harmonicity and Beating Cues. *Journal of Cognitive Neuroscience*, 2023, 35 (5), pp.765-780. 10.1162/jocn_a_01973 . hal-04080670

HAL Id: hal-04080670

<https://hal.science/hal-04080670>

Submitted on 25 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Consonance perception in congenital amusia:
behavioral and brain responses to harmonicity and beating cues

Jackson E. Graves^{1,2,3,a}, Agathe Pralus¹, Lesly Fornoni¹, Andrew J. Oxenham³, Barbara
Tillmann^{1*}, and Anne Caclin^{1*}

*Joint last authors

¹Lyon Neuroscience Research Center, CNRS UMR 5292, Inserm U1028, Université Lyon 1,
Lyon, France

²Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale
supérieure, PSL University, CNRS, Paris, France

³Department of Psychology, University of Minnesota, Minneapolis, MN, USA

^a Author to whom correspondence should be addressed.
Electronic mail: jackson.graves@ens.psl.eu

Abstract

Congenital amusia is a neurodevelopmental disorder characterized by difficulties in the perception and production of music, including the perception of consonance and dissonance, or the judgment of certain combinations of pitches as more pleasant than others. Two perceptual cues for dissonance are inharmonicity (the lack of a common fundamental frequency between components) and beating (amplitude fluctuations produced by close, interacting frequency components). Amusic individuals have previously been reported to be insensitive to inharmonicity, but to exhibit normal sensitivity to beats. In the present study, we measured adaptive discrimination thresholds in amusic participants and found elevated thresholds for both cues. We recorded EEG and measured the mismatch negativity (MMN) in evoked potentials to consonance and dissonance deviants in an oddball paradigm. The amplitude of the MMN response was similar overall for amusic and control participants; however, in controls there was a tendency towards larger MMNs for inharmonicity than for beating cues, whereas the opposite tendency was observed for the amusic participants. These findings suggest that initial encoding of consonance cues may be intact in amusia despite impaired behavioral performance, but that the relative weight of non-spectral (beating) cues may be increased for amusic individuals.

Keywords: amusia; consonance; MMN; inharmonicity; beating

Consonance perception in congenital amusia:
behavioral and brain responses to harmonicity and beating cues

1. INTRODUCTION

A fundamental aspect of Western tonal music is the phenomenon of consonance and dissonance, whereby certain combinations of simultaneous pitches are judged to be more pleasant, and to create less tension, than others. The perception of consonance and dissonance is determined at least in part by enculturation (Butler & Daston, 1968; Lahdelma & Eerola, 2020; Lundin, 1947; McDermott et al., 2016; Popescu et al., 2019), but our ability to discriminate consonant from dissonant sounds is likely to reflect differences in the acoustic properties of those sounds. Listeners from Western cultures tend to rate pitch intervals with simple frequency ratios as more consonant than intervals with complex ratios, a phenomenon with at least two potential explanations. Dissonant (i.e., complex ratio) intervals generally produce *beating*: upper harmonics of the two tones can be close together in frequency and so interact in the auditory periphery, leading to amplitude fluctuations that are perceived as roughness, which is often described as unpleasant (Hutchinson & Knopoff, 1978; Plomp & Levelt, 1965; von Helmholtz, 1912). Dissonant intervals also lack *harmonicity*: their components cannot be interpreted as multiples of a single common fundamental frequency (Ebeling, 2008; Stumpf, 1883; Terhardt, 1974). One study testing the relative importance of these two cues in young Western listeners found that inharmonicity was more likely than beating to underlie the perception of dissonance, particularly in those with some formal musical training (McDermott et al., 2010).

Further insight into consonance perception comes from the study of congenital amusia, a disorder characterized by difficulties in the perception and production of music in individuals

with clinically normal hearing and no general cognitive deficits (Ayotte et al., 2002), which affects 2-4% of the population (Kalmus & Fry, 1980; Peretz & Vuvan, 2017). Deficits exhibited by individuals with amusia can provide insights about general mechanisms of music perception, and of auditory perception more generally. Pleasantness ratings of chords have been found to be impaired to some degree in congenital amusia, with control participants rating some chords (e.g., a perfect fifth) as more pleasant than others (e.g., an augmented fourth) and amusic participants showing less pronounced distinctions (Cousineau et al., 2012; Marin et al., 2015). One of these previous studies tested the perception of consonance and dissonance cues in amusic participants and found that they exhibited behavioral impairment for the detection of inharmonicity, but not for the detection of beating (Cousineau et al., 2012), lending further support to the view that consonance perception is dominated by harmonicity, at least in Western musical cultures (McDermott et al., 2010).

Congenital amusia is generally thought of as a disorder affecting pitch perception (Cousineau et al., 2015; Peretz, 2016; Tillmann et al., 2015) and the perception of timbral cues relating to the frequency content of sounds (Graves et al., 2019; Marin et al., 2012), leaving intact other capabilities, such as memory for speech stimuli (Albouy, Peretz, et al., 2019; Tillmann et al., 2009), and with more restricted or conditional deficits for other musical attributes, like rhythm (Foxton et al., 2006; Hyde & Peretz, 2004) and loudness (Graves et al., 2019; Tillmann et al., 2016). However, a recent study (Whiteford & Oxenham, 2017) found that amusic individuals were also impaired at detecting amplitude modulation (AM) as well as frequency modulation (FM), suggesting that the deficit may not be limited to cues related to pitch and timbre. The finding that amusics are impaired at AM detection has interesting implications for amusic perception of beating, and seems to conflict with the previously reported

lack of amusic impairment for beating perception (Cousineau et al., 2012). This discrepancy may be due to the different levels of AM tested by the two studies: Cousineau et al. (2012) used stimuli that produced the equivalent of 100% depth of AM (well above threshold), whereas Whiteford and Oxenham (2017) measured pure-tone AM detection abilities at modulation depths near detection threshold. However, thresholds for the detection of beating in a complex tone have not been directly measured for amusic individuals, and further exploration of consonance perception in amusia is warranted to resolve this apparent discrepancy.

Studies of brain function and anatomy in amusia generally point to an origin of the deficit in the network connecting auditory areas in temporal cortex with higher-level areas in frontal cortex (Albouy et al., 2013, 2015; Hyde et al., 2006, 2007, 2011; Leveque et al., 2016). In terms of early cortical encoding of pitch, fMRI studies have not observed differences in amusics' activation patterns with classical subtraction analyses (Hyde et al., 2011; Norman-Haignere et al., 2016). However, a re-analysis with whole-brain multivariate pattern analyses revealed that amusic and control participants differed in their pattern of functional activity in right Heschl's gyrus (Albouy, Caclin, et al., 2019). In addition, an MEG study observed differences in amusics and controls at the early level of the N100m response (Albouy et al., 2013), leaving open the possibility that early pitch processing may also be affected in amusia.

Automatic sound processing in amusia has previously been studied using the mismatch negativity (MMN), an EEG potential recorded in response to sounds that violate low-level auditory expectations, and that can occur even in the absence of directed attention (Näätänen et al., 1978, 2007). Despite showing behavioral impairment at pitch discrimination tasks, amusic individuals have been found to exhibit normal MMN responses, even to small pitch changes that are below their behavioral discrimination threshold (Moreau et al., 2009, 2013; Peretz et al.,

2009). A similar pattern of overall normal MMNs in amusia has also recently been observed for emotional prosody (Pralus et al., 2020) and in pitch sequences of various levels of melodic complexity (Quiroga-Martinez et al., 2021), yet with differences in an early negativity preceding the MMN as well as in the P3a (Pralus et al., 2020). One EEG study found a reduced early frontal negativity in response to unpredictable notes in a melody (Omgie et al., 2013), suggesting that potential impairments of automatic brain responses at this early level may still exist for amusics for certain perceptual cues.

For normal-hearing, non-amusic listeners, the MMN has also been reported in response to harmonicity changes and to pitch differences of harmonic complex tones (Butler & Trainor, 2012; Tervaniemi, Schröger, Saher, & Näätänen, 2000). In addition to high-level cues like tonal expectation (Marmel et al., 2011), the MMN has been observed in response to differences in consonance and dissonance (Brattico et al., 2009; Crespo-Bojorque et al., 2018; Linnavalli et al., 2020). More specifically, an MMN has been measured in response to inharmonicity in the form of a mistuned harmonic (Alain et al., 2001), to harmonicity in the form of a harmonic complex presented in the context of inharmonic complexes (Jones, 2003), and to roughness cues created by beating-induced AM (De Baene et al., 2004). The MMN may therefore be a good candidate for indexing early neural processing of consonance and dissonance in amusia, as its presence has been established in response to pitch changes for amusic participants, and in response to dissonance cues for normal-hearing listeners.

In the present study, we measured behavioral sensitivity and MMN for amusic and control participants to two potential cues for consonance and dissonance: inharmonicity (harmonic complex tones vs. linearly shifted harmonics, resulting in an inharmonic series) and beating (inharmonic complex tones with or without spectral sideband components added to

produce beats). In the first experiment, participants' detection thresholds were measured for both inharmonicity and beating cues, as well as for pure-tone AM to assess basic sensitivity to amplitude fluctuations (Whiteford & Oxenham, 2017). In the second experiment, participants listened either actively or passively to streams of stimuli in an oddball paradigm, where a dissonant stimulus deviated from a standard consonant stimulus, or vice versa. If behavioral impairments for consonance perception in amusia (Cousineau et al., 2012; Marin et al., 2015) can be traced to differences in early encoding of these cues, this could be reflected in reduced MMN amplitude for amusic individuals. If, on the other hand, the deficits arise from later processes, amusic individuals should show normal MMN responses to these stimuli.

2. EXPERIMENT 1: BEHAVIORAL DISCRIMINATION

2.1. Methods

2.1.1. Participants

Amusia is rare in the general population (Peretz & Vuvan, 2017), making recruitment challenging. However, the effect sizes for relevant behavioral measures, such as pitch discrimination or inharmonicity detection (e.g. Cousineau et al., 2012), have been very large when comparing amusic and control groups. For instance, the partial eta-squared (η_p^2) values for pitch discrimination thresholds reported by Whiteford and Oxenham (2017) were greater than 0.3, more than twice the value generally considered a large effect size ($\eta_p^2 = 0.14$; Cohen, 1988). Assuming an effect size of $\eta_p^2 = 0.14$ for a within-between interaction effect in an ANOVA, the sample size required to detect this effect with 80% power is only 8 participants per group. Similarly, for inharmonicity detection, an examination of the individual data reported by Cousineau et al. (2012; their Fig. 6) suggests a Cohen's d value of 3.28, which is more than four

times the value generally considered a large effect ($d = 0.8$; Cohen, 1988). Assuming an effect size of at least $d = 1.5$ in the current study, the sample size required to detect this effect with 80% power is only 9 participants per group. Most previous behavioral studies of amusia have recruited at least 10-15 participants per group (e.g. Cousineau et al., 2012; Whiteford & Oxenham, 2017). We therefore recruited 12 amusic individuals (6 female and 6 male) and 11 control participants (7 female and 4 male) for our study measuring behavioral thresholds.

During recruitment, control participants were matched to amusic participants in terms of age, education, gender, and handedness (see Table 1). No participants had any formal musical training, except for one control participant who reported one year of musical training. All participants had normal hearing as determined by pure tone audiometry, and congenital amusia was identified with the Montreal Battery of Evaluation of Amusia (MBEA) (Peretz et al., 2003). Following the MBEA, each participant's frequency difference limen (FDL) was also measured using the adaptive tracking procedure described in Tillmann et al. (2009) for a pure tone at 512 Hz. All participants provided written informed consent and were compensated for their participation. Study procedures were approved by the required ethics committee, the Comité de Protection des Personnes (CPP), Sud-Est II.

	Controls (N=11)	Amusics (N=12)	<i>p</i> value
Age (years)	37.0 (16.3)	38.7 (14.9)	0.79
Education (years)	15.7 (2.1)	15.4 (2.8)	0.76
FDL (semitones)	0.20 (0.13)	2.08 (1.56)	< 0.001
MBEA: global (percent)	89.3 (3.0)	71.3 (6.5)	< 0.001
MBEA: pitch (percent)	89.4 (5.2)	68.9 (7.8)	< 0.001

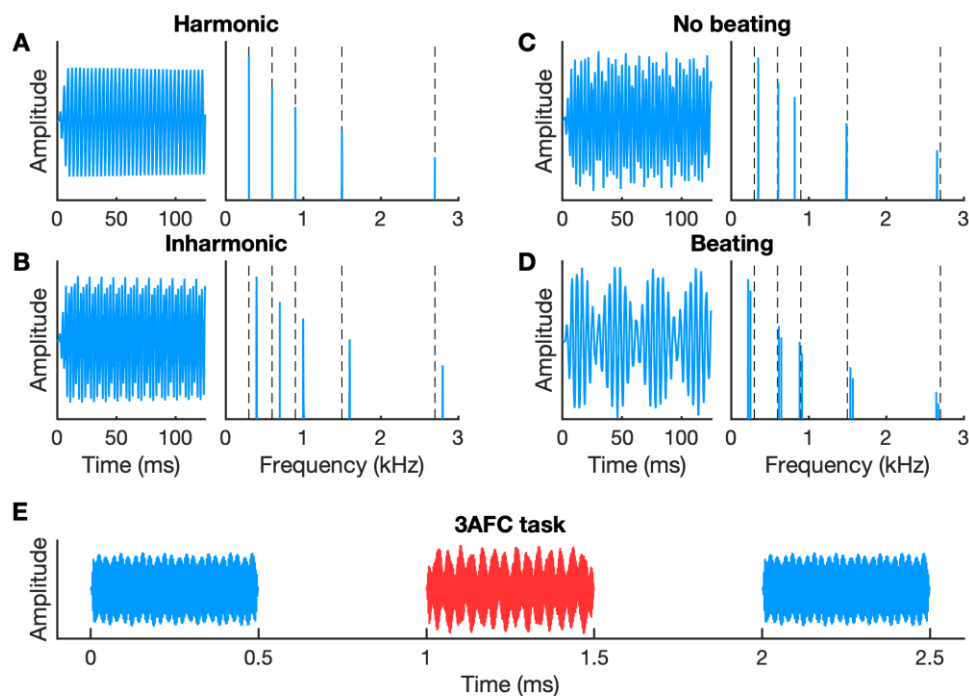
Table 1. Comparison of amusic and control samples in Experiment 1 by demographic variables and music perception scores. Mean values are listed, with standard deviations in parentheses. Bold indicates a p value less than 0.05 in a two-tailed independent t-test between groups. Congenital amusia is identified using the Montreal Battery of Evaluation of Amusia (MBEA), with a global cutoff score of 78% correct over six subtests (Peretz et al., 2003). MBEA scores for the pitch-related subtests only (Scale, Interval, and Contour) are also listed, along with frequency difference limens (FDL) measured using an adaptive tracking procedure (Tillmann et al. 2009).

2.1.2. Stimuli

Complex tones, inspired by the stimuli of Cousineau et al. (2012), were generated in four categories to manipulate consonance cues: harmonic, inharmonic, no-beating, and beating (see Figure 1). Harmonic complexes consisted of the 1st, 2nd, 3rd, 5th, and 9th components of a harmonic series, added together in sine phase. Inharmonic complexes were generated by shifting every component in the harmonic complex up by a uniform distance in linear frequency. When manipulating beating cues, we wanted to be sure that the perception of dissonance was not due to inharmonicity from the introduction of non-harmonic components, but purely due to beating. The stimuli were therefore always inharmonic, with frequency components chosen by sampling from a uniform distribution between -30 and +30 Hz relative to each component in the equivalent harmonic stimulus, added together in random phase. Beating stimuli were created by adding a single sideband component 30 Hz away from each component, with a uniform level difference between sidebands and original components. The location of sidebands (above or below original components) was randomly chosen for each complex but was consistent within the complex. See supplemental audio files SA1 - SA4 for examples of complexes in each of the four categories^b.

^b Supplemental material can be found at: <https://osf.io/zmeaq/>

Figure 1. Stimuli manipulating harmonicity and beating cues. Temporal waveforms and frequency spectra of stimuli for harmonicity (left) and beating (right) conditions, with vertical dashed lines showing the 1st, 2nd, 3rd, 5th, and 9th components in a harmonic series. Harmonic complexes (A) contained the five components shown, with a decaying spectral envelope. Inharmonic complexes (B) were shifted up relative to harmonic complexes by a uniform distance in linear frequency. No-beating stimuli (C) were inharmonic, with frequency components jittered relative to a harmonic series. Beating stimuli (D) introduced a single sideband component 30 Hz above or below each component in the no-beating stimuli, with a uniform level difference between sidebands and original components. (E) shows the timing of a trial on the 3AFC discrimination task used in Experiment 1. Participants identified the target (red) by comparing against two references (blue), with target position randomized on each trial



In addition to our manipulation of beating, which introduced AM to a complex tone, we also directly measured thresholds for detection of pure-tone sinusoidal AM. Stimuli for the AM detection test involved a pure-tone carrier with a frequency that was varied between 277 and 370 Hz. When AM was present, the modulation frequency was constant at 10 Hz. This was chosen to be at an intermediate frequency from the ones tested by Whiteford & Oxenham (2017) of 4 and 20 Hz. Sinusoidal AM (SAM) detection differs from beat detection because SAM involves two symmetrical placed spectral sidebands, whereas beat detection involves only one; however, SAM and beat thresholds are closely related and are independent of modulation rate between 10 and 30 Hz (Kohlrausch et al., 2000).

To ensure that the perception of dissonance cues was based on the dissonance itself and not on a perceived difference in pitch, we randomly varied the fundamental frequency (F0) between intervals of all tones used in both experiments. The F0 of each complex tone (or carrier frequency for AM stimuli, or nominal F0 for all inharmonic complex tones) was chosen from among six discrete values, at semitone increments between C#4 (277 Hz) and F#4 (370 Hz). F0s were chosen such that throughout both experiments, no two consecutive tones shared the same F0, including both standard and deviant tones. All complex tones had a decaying spectral envelope, such that components decreased in level at a rate of 14 dB/octave from the lowest frequency component. The duration of each tone was 500 ms, including 10-ms raised-cosine onset and offset ramps, and the overall level of each tone was 65 dB sound pressure level (SPL). All stimuli were created using MATLAB (The Mathworks, Natick, MA) and presented diotically at a sampling rate of 48 kHz through circumaural Sennheiser HD 280 pro headphones in a quiet sound booth.

2.1.3. Procedure

Participants discriminated inharmonic from harmonic, beating from no-beating, or AM from no-AM stimuli in a 3AFC task. On each trial, participants heard three tones, and were asked to select the tone that sounded different from the other two. The target tone was always the more dissonant sound (inharmonic, beating, or AM), while the two reference tones were more consonant (harmonic, no beating, or no AM). The size of the spectral shift for inharmonic stimuli and the amplitudes of the sidebands for beating stimuli were adaptively varied in order to measure threshold. During this adaptive procedure, the size of the inharmonic frequency shift varied on a logarithmic scale between 1 and 125 Hz and the amplitude of the sidebands varied between -60 and 0 dB relative to the original components. Listeners were instructed to ignore the random changes in F0 and to identify the tone that stood out from the other two (without explaining explicitly on which feature). Feedback on whether the response was correct or incorrect was provided immediately after each trial.

The difference between target and reference was adaptively varied, and discrimination thresholds were measured using a tracking procedure following a 2-down 1-up rule that tracks the 70.7% point on the psychometric function (Levitt, 1971). An adaptive run began at 50 Hz and changed by factors of 2.51, 1.58, and 1.26 (inharmonic shift) or began at -10 dB and changed by step sizes of 4, 2, and 1 dB (sideband and AM depth). The step size changed every two reversals until the final step size, after which the run continued for a further six reversals at the final step size. Each participant completed three runs for each cue (harmonicity, beating, and AM) and the values at the six last reversals from each run were averaged (geometric means for inharmonicity) for an estimated threshold. The order of conditions was partially counterbalanced, with half of participants performing the tasks in the order of harmonicity, AM, and beating, and

the other half performing the tasks in the order of AM, beating, and harmonicity. Each condition was preceded by a brief orientation period, consisting of three labeled example trials with large differences equal to the starting value.

2.2. Results

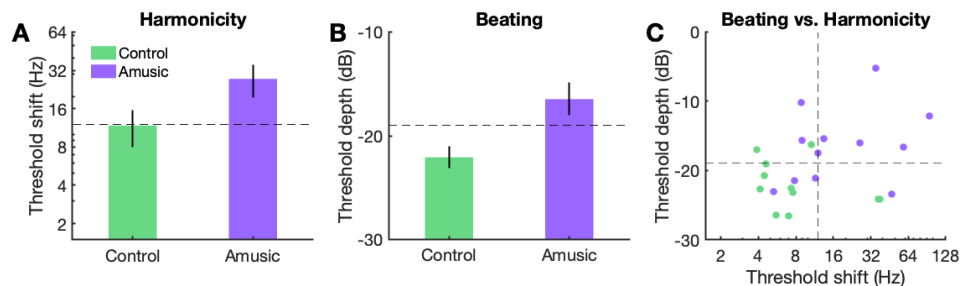


Figure 2. Behavioral results for discrimination of harmonicity and beating cues by amusic and control participants. Top: thresholds measured with 3AFC adaptive tracking (Experiment 1) for inharmonic shift (harmonicity, A) and single sideband depth (beating, B), for controls (green, $N = 11$) and amusics (purple, $N = 12$). Dashed lines indicate levels chosen for the constant-stimuli oddball paradigm in the EEG study. Error bars show ± 1 SEM. Individual thresholds are shown in (C).

Inharmonic shift thresholds were larger for amusic participants than for control participants [$t(1,21) = 2.24$, $p = 0.036$, $d = 0.94$] (see Figure 2A). A similar effect was observed for both of the beating-related tasks: complex single-sideband beating detection [$t(1,21) = 2.91$, $p = 0.0084$, $d = 1.23$] (Figure 2B) and pure-tone SAM detection [$t(1,21) = 2.98$, $p = 0.0071$, $d = 1.27$] (Figure S1). Although the deficits associated with amusia are often considered to be pitch specific, here amusic participants exhibited reduced behavioral sensitivity to non-pitch cues in beating and pure-tone SAM detection. The similar effect sizes suggest that the deficit in detecting beating may be as large under amusia as the deficit in detecting inharmonicity. As expected, detection of beating in a complex tone was comparable to detection of pure-tone SAM. In addition, thresholds for these two tasks correlated with each other [$r = 0.73$, $p < 0.001$], and

this correlation was significant within the amusic group [$r = 0.68$, $p = 0.015$], but not the control group [$r = 0.41$, $p = 0.21$]. A correlation was also observed between SAM detection and inharmonicity detection [$r = 0.56$, $p = 0.005$] (Figure S1), again significant within the amusic group [$r = 0.58$, $p = 0.049$], but not within the control group [$r = -0.24$, $p = 0.48$]. No significant correlation was observed between thresholds for beating and inharmonicity, whether across groups [$r = 0.30$, $p = 0.16$], within amusics [$r = 0.26$, $p = 0.42$], or within controls [$r = -0.27$, $p = 0.41$] (Figure 2C).

3. EXPERIMENT 2: EEG MEASUREMENT OF THE MISMATCH NEGATIVITY

3.1. Methods

3.1.1. Participants

Most previous EEG studies of amusia have recruited 10-20 participants per group (e.g. Moreau et al., 2013; Quiroga-Martinez et al., 2021), so for our EEG study, we recruited 19 amusic individuals (10 female and 9 male) and 21 control participants (13 female and 8 male). Out of the participants in Experiment 2, 13 individuals (5 amusics and 8 controls) had also participated in Experiment 1. Participants in Experiment 2 took part in two consecutive MMN studies within the same session; the first one with emotional vowels is reported in Pralus et al. (2020) and the second one is reported here.

	Controls (N=21)	Amusics (N=19)	<i>p</i> value
Age (years)	32.3 (14.4)	30.7 (14.4)	0.73
Education (years)	15.3 (2.2)	15.0 (2.7)	0.71
FDL (semitones)	0.29 (0.15)	1.33 (1.48)	0.003
MBEA: global (percent)	88.2 (3.5)	73.4 (6.0)	< 0.001
MBEA: pitch (percent)	88.7 (4.8)	70.2 (6.6)	< 0.001

As in Experiment 1, control participants were matched to amusic participants in terms of age, education, gender, and handedness (see Table 2). The only participant who reported any musical training was one control participant who reported one year of musical training (not the

Table 2. Comparison of amusic and control samples in Experiment 2 by demographic variables and music perception scores. Mean values are listed, with standard deviations in parentheses. Bold indicates a p value less than 0.05 in a two-tailed independent t-test between groups. Details as in Table 1.

same participant as in Experiment 1). Congenital amusia was identified with the MBEA, and FDLs were measured with the same procedure as in Experiment 1. All participants provided written informed consent and were compensated for their participation. Study procedures were approved by the required ethics committee, the Comité de Protection des Personnes (CPP), Sud-Est II.

3.1.2. Stimuli

All tones used in Experiment 2 were generated according to the same rules as in Experiment 1, in one of four types: harmonic, inharmonic, no-beating or beating (see Figure 1 and Experiment 1 stimuli). For Experiment 2, a constant shift size and sideband amplitude were chosen to differentiate the standard and deviant stimuli to be played to every participant. These levels were chosen to maximize the difference in discriminability between participant groups (amusic and control) observed in Experiment 1. Accordingly, we chose an inharmonic shift of 12 Hz, which was above threshold for 82% of control participants, but was above threshold for only 42% of amusic participants, and a sideband amplitude of -19 dB (relative to the original tones), which was above threshold for 82% of control participants, but was above threshold for only

33% of amusic participants. Auditory stimuli were presented diotically via MCL Samar CSQ-ECM insert earphones during EEG recording.

3.1.3. Procedure

During EEG recording, streams of tones were presented to participants with a stimulus onset asynchrony (SOA) of 700 ms. In each of four conditions, one stimulus type was defined as standard, occurring on 5/6 of trials, and its opposite as deviant, occurring on only 1/6 of trials (see Figure 3). Conditions were named after the deviant tone: harmonic-deviant (HDEV), inharmonic-deviant (IDEV), no-beating-deviant (NDEV), and beating-deviant (BDEV). Deviants were always separated from each other in the stream by no more than 7 and no fewer than 3 standards. The F0 of each tone was selected randomly, with the constraint that no two consecutive tones shared the same F0. Each participant completed four blocks of passive listening followed by four blocks of an active task (see below). Passive blocks always came before active blocks. The order of the four blocks was always the same in passive and active phases, but the order of blocks was counterbalanced across participants. Passive blocks contained 140 deviants and 700 standards in total, and participants were instructed to sit still and watch a silent subtitled movie. Active blocks contained 30 deviants and 150 standards, and participants were instructed to press a button every time they heard a sound that was different from most other sounds.

As a measure of behavioral sensitivity to the difference between standard and deviant, the sensitivity index d' was calculated from active-task responses in the following manner: Any responses between 100 and 2200 ms after a deviant onset were counted as a hit for that deviant, and any responses between 2200 ms after one deviant onset and 100 ms after the following deviant onset were counted as a false alarm. In this way, the duration of hit periods was constant

at 2100 ms per deviant, while the duration of false alarm periods was variable from 700 to 3500 ms, but averaged 2100 ms, meaning that a random responder would make roughly equal numbers of hits and false alarms. The value of d' was then calculated using the resulting hit and false alarm rates (Stanislaw & Todorov, 1999).

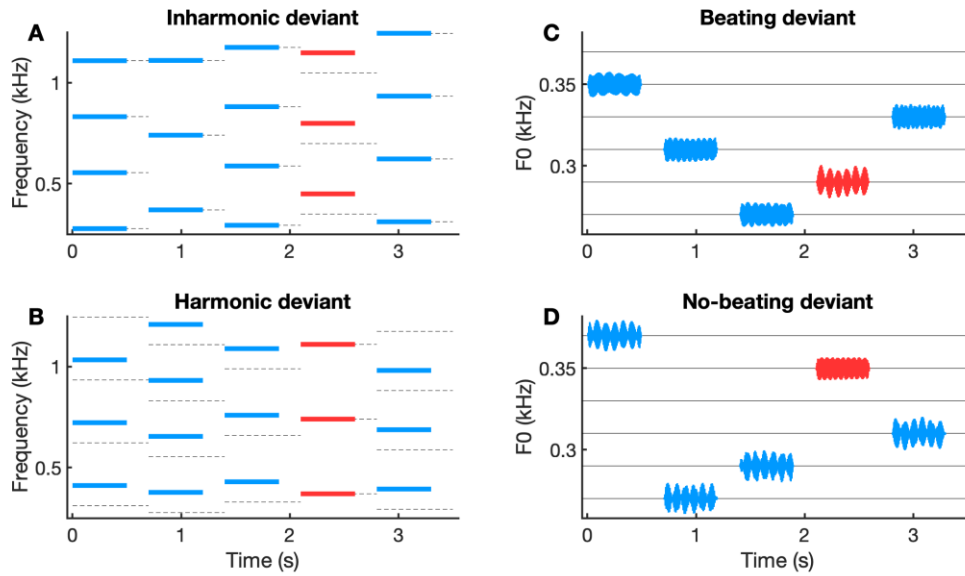


Figure 3. Oddball paradigm with consonance changes and pitch roving. Participants heard deviant (red) and standard (blue) stimuli in four different experimental conditions, while nominal F0 varied constantly, with a minimum distance of one semitone between each pair of two consecutive tones. Each of the four stimulus types (harmonic, inharmonic, no-beating and beating) served as the deviant in one condition, with its contrast serving as the standard. Left panels: sound spectra for inharmonic-deviant (A) and harmonic-deviant (B) conditions; dashed grey lines indicate harmonic partials. Right panels: sound waveforms for beating-deviant (C) and no-beating-deviant (D) conditions, plotted at the level of their fundamental frequency (F0). In the beating-deviant and no-beating-deviant conditions, all sounds were inharmonic.

3.1.4. EEG recording and preprocessing

EEG was recorded from 31 Ag/AgCl active electrodes (Acticap, Brain Products GmbH, Gilching, Germany), with a nose reference, at a sampling rate of 1000 Hz with a 0.016-1000 Hz bandwidth (BrainAmp Standard Amplifier, Brain Products GmbH, Gilching, Germany). Vertical eye movements were recorded with an additional active electrode positioned under the left eye (offline re-referenced to Fp1). Pre-processing was conducted in MATLAB using functions from eeglab (Delorme & Makeig, 2004). Data was bandstop filtered to remove line noise, using 4th-order Butterworth filters with cutoffs of 47 and 53 Hz, then 147 and 153 Hz. A 20-s portion of data containing eyeblinks but few other artifacts was visually selected for each participant, for the purpose of performing an independent components analysis (ICA). ICA weights from this subset of data were visually examined, eyeblink components were selected and then removed from the entirety of the data through an ICA inverse transformation. After ICA-based blink correction, the data were divided into 700-ms epochs, from -200 ms to +500 ms relative to stimulus onset, and non-blink artifacts were rejected using a dynamic range threshold applied to each epoch. This threshold, and the channels to be ignored for artifact rejection, were manually selected for each participant based on visual inspection of the data (thresholds were constant across channels for each participant, and ranged from 90 μ V to 300 μ V, with an average of 153 μ V; the number of excluded channels ranged from 0 to 21, with an average of 7.59). Of the 140

deviant trials in each passive block, the average number rejected due to artifacts was 31, leaving an average of 109 deviants for analysis per passive block. After artifact rejection, epochs were averaged for each condition, separately for deviants and standards, excluding standards immediately following a deviant. The resulting ERPs were bandpass filtered (Butterworth, 4th order) from 2 to 30 Hz, and the baseline from 100 ms to 0ms before stimulus onset was subtracted. Finally, any exceptionally noisy electrodes were removed and re-interpolated from surrounding electrodes based on their three-dimensional coordinates.

3.1.5. Analysis of EEG data

Only EEG data from the passive listening blocks were analyzed. The MMN was evaluated by subtracting ERPs to standard sounds from their acoustically identical deviant sounds. This means, for example, that the harmonic deviant (from the HDEV block) was compared against the harmonic standard (from the IDEV block), not against the (inharmonic) standard from its own block. In this way, any measured difference between ERPs cannot be due to acoustic differences between the sounds, as they are the same sound, and is likely instead to arise from neural encoding of the deviant sound as deviant. For each deviant-standard pair, ERP amplitudes were evaluated for the whole brain over pre-defined periods of interest to assess whether significant MMN and/or P3a emerge (see below). Spatiotemporal clusters of significant differences between standard and deviant ERP amplitudes over each period were identified using non-parametric cluster-based permutation statistics (Blair & Karniski, 1993), implemented with the Fieldtrip toolbox in MATLAB (Maris & Oostenveld, 2007). Using this method, contiguous spatiotemporal clusters of differences between standard and deviant were identified, and the significance of a cluster was determined by comparing the sum of t-values in this cluster to the sum of t-values in the largest cluster identified when standard and deviant labels were randomly

permuted for each subject over 10,000 iterations. The associated cluster-level p -value was the proportion of random permutations that resulted in a cluster larger than the one identified.

We used cluster-based permutation statistics to identify electrode sites and temporal windows where a significant MMN emerged (defined for our study as a negativity within the 150-300 ms time range) or where a significant P3a emerged (for our study, a positivity in the 300-450 ms time range). These clusters were identified across all participants, separately for each of the four deviant types (harmonic, inharmonic, no-beating, and beating). The difference values (deviant minus standard) within each spatiotemporal cluster were then averaged together for comparison by group (amusic or control), feature (beating or harmonicity) and consonance (consonant or dissonant).

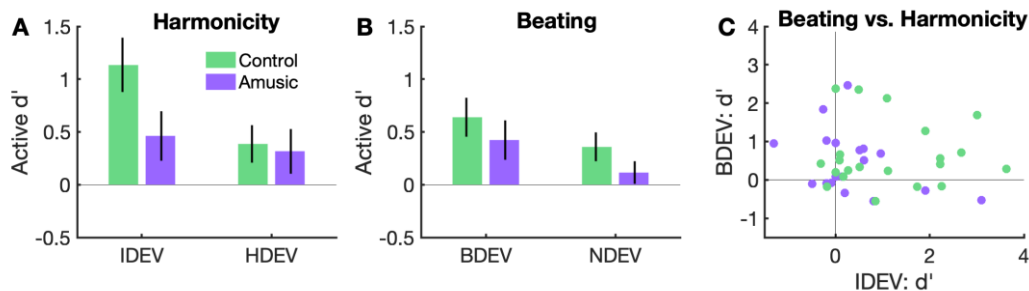
3.2. Results

3.2.1. *Decreased behavioral sensitivity to harmonicity and beating changes in amusia*

When tested behaviorally during the active task blocks of Experiment 2, deviant detection was impaired in amusic participants for both cues (see Figure 4). We conducted a nonparametric^c version of a 3-way mixed ANOVA on sensitivity (d') considering the between-subjects factor of participant group (amusic or control) and within-subjects factors of feature (harmonicity or beating) and consonance (consonant vs. dissonant deviants), using the nparLD package in R (Noguchi et al., 2012). We found a main effect of consonance [$F(1,38) = 10.61, p = 0.0011, \eta_p^2 = 0.32$], reflecting better performance on dissonant-deviant blocks (IDEV and BDEV, as measured adaptively in Experiment 1) than on consonant-deviant blocks (HDEV and NDEV). We also found a main effect of participant group [$F(1,38) = 4.74, p = 0.029, \eta_p^2 = 0.09$], reflecting better performance for the control than the amusic group. No other significant main effects or interactions were found (all $p > 0.35$).

^c Because d' is not normally distributed, with a mode at 0 (representing chance performance), but a longer positive tail (representing better-than-chance performance), a nonparametric test is more appropriate.

Figure 4. Behavioral results during the active task of Experiment 2. Behavioral sensitivity (d') for identifying deviants is plotted for the harmonicity contrast (A) with an inharmonic deviant (IDEV) or a harmonic deviant (HDEV), and for the beating contrast (B) with a beating deviant (BDEV) or a no-beating deviant (NDEV), for control (N = 21) and amusic (N = 19) groups. Chance performance corresponds to a d' of 0. Individual performance for IDEV and BDEV is shown in (C).



Some participants had previously participated in the threshold measurement study, in which the 3AFC task with its dissonant target resembled the dissonant-deviant conditions more than the consonant-deviant conditions. In order to test whether the apparent effect of consonance was merely an effect of practice from Experiment 1, we re-ran the ANOVA excluding the 13 participants who had completed Experiment 1. The main effect of consonance remained after removing these subjects [$F(1,25) = 5.09$, $p = 0.024$, $\eta_p^2 = 0.27$], suggesting that this effect was not due to participants having practiced, and may reflect an inherent perceptual asymmetry in the stimuli, perhaps due to processes of auditory feature extraction (Cusack & Carlyon, 2003).

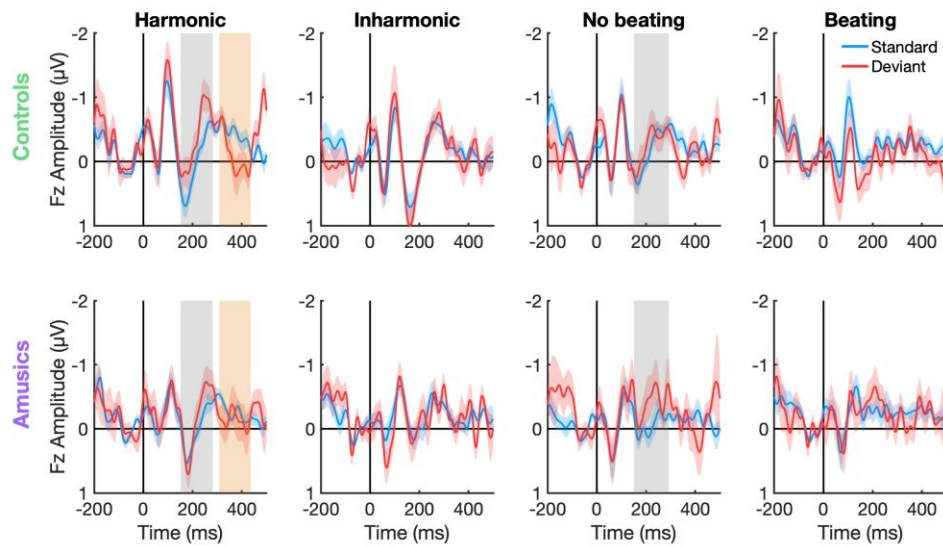
Participant performance was evaluated against chance with Wilcoxon signed rank tests of d' against zero. Control participants performed significantly better than chance in all four

conditions (IDEV: $p = 0.0034$, HDEV: $p = 0.011$, BDEV: $p = 0.0065$, NDEV: $p = 0.016$), while amusic participants only performed better than chance in the IDEV condition ($p = 0.048$) and the BDEV condition ($p = 0.036$). If Bonferroni multiple-comparisons correction was applied ($\alpha = 0.05 / 8 = 0.00625$), only control participants performed better than chance, and only in the IDEV condition. These results, along with the significant main effect of group on d' , suggest that behavioral discrimination of the stimuli used in the EEG study was more difficult for the amusic than for the control group, as expected based on the results from Experiment 1.

3.2.2. Presence of the MMN and P3a in response to consonant sounds

Each deviant ERP was compared against its acoustically identical standard (see Figure 5). In general, peak negativity of deviants relative to standards occurred in frontal electrodes between 150 and 300 ms, followed by positivity in frontal electrodes between 300 and 450 ms (see Figures 6 and S2). Significant MMNs were observed only for conditions HDEV and NDEV.

Figure 5. ERPs to standard and deviant stimuli at Fz by group and condition. Responses are shown at electrode Fz for each of the four types of stimulus, for controls (top row) and amusics (bottom row). Each plot compares a deviant to its acoustically identical standard (the same sound in a different context from the opposite condition). Blue and red colored regions show ± 1 SEM. Gray regions indicate significant MMN across all participants, observed for harmonic and no-beating sounds. Orange regions indicate significant P3a across all participants, observed for harmonic sounds.



Across all participants, significant MMN clusters were identified for harmonic deviants compared to harmonic standards ($p = 0.034$, 150-290 ms) and no-beating deviants compared to no-beating standards ($p = 0.011$, 150-300 ms), but not for inharmonic or beating stimuli (cluster-

level $p > 0.73$). See Figure S3 for the temporal course of MMN topographies. A significant P3a cluster was identified for harmonic stimuli ($p = 0.026$, 300-440 ms), but not for the other three sounds (cluster-level $p > 0.48$).

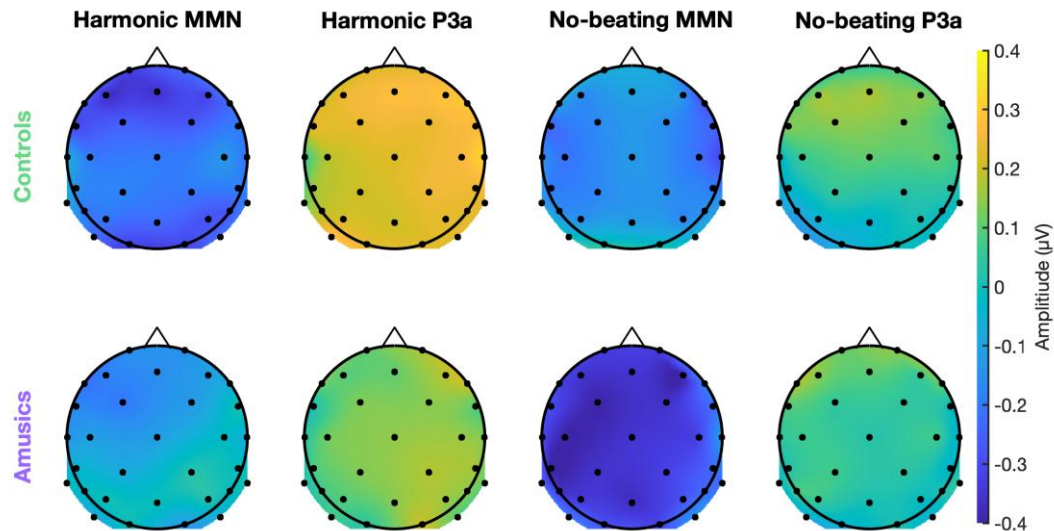


Figure 6. Topography of MMN and P3a responses to consonant stimuli. The difference amplitude (deviant minus standard) across the scalp is shown, averaged over two time periods: 150-300 ms (MMN, odd columns) and 300-450 ms (P3a, even columns), for harmonic sounds (left) and no-beating sounds (right), for controls (top) and amusics (bottom). Small black dots indicate electrode positions. See Figure S2 for the topographies of non-significant responses for dissonant deviants.

3.2.3. Amplitude of the MMN by consonance, group, and feature

We compared the amplitude of ERPs across participant groups, features, and consonance levels by running a 3-way mixed ANOVA on MMN magnitudes, considering the factors of consonance (consonant vs. dissonant deviant), feature (beating vs. inharmonicity), and group (amusic vs. control). It is important to note that conditions in this ANOVA differed from each other based on the contrast between deviant and standard, and not based on the deviant or standard alone. MMNs were averaged within the relevant spatiotemporal cluster: for harmonic and inharmonic stimuli, the harmonic MMN cluster; for no-beating and beating stimuli, the no-

beating MMN cluster (see Figure 7). We observed a main effect of consonance on MMN amplitude [$F(1,38) = 15.13, p < 0.001, \eta_p^2 = 0.28$], reflecting larger MMNs for harmonic and non-beating stimuli than for inharmonic and beating stimuli, in keeping with the fact that the presence of the MMN was only confirmed in the HDEV and NDEV conditions (see section 3.2.2). This effect ran opposite to the effect of consonance on behavior, as stronger MMNs were observed in response to consonant deviants, but better behavioral performance was found for dissonant deviants. As with the behavioral results, we re-ran this analysis without the 13 subjects who had completed Experiment 1, in order to check whether it might be related to practice on the 3AFC threshold task with dissonant targets. After removing these subjects, the effect of consonance on MMN strength persisted [$F(1,25) = 13.18, p = 0.0013, \eta_p^2 = 0.35$], suggesting it did not depend on practice in the threshold task.

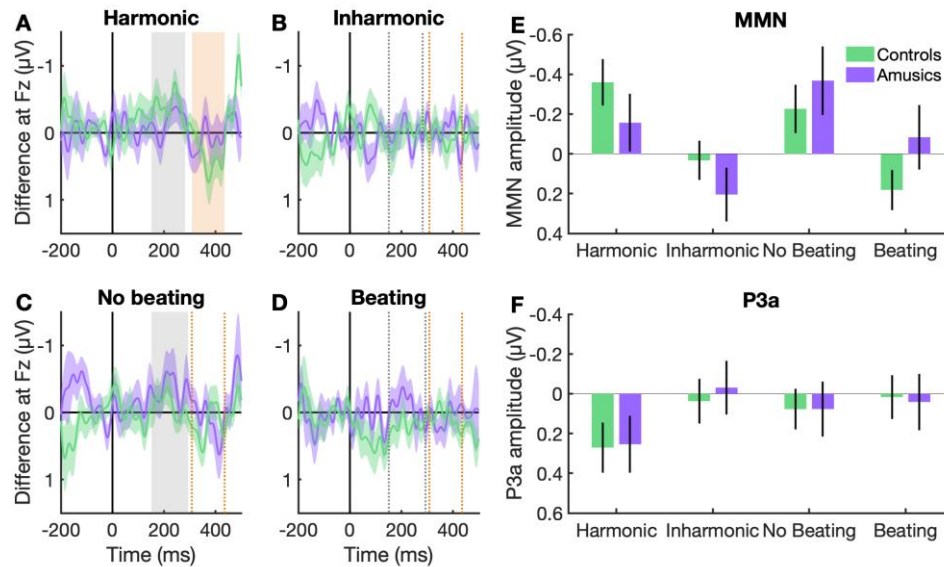


Figure 7. MMN amplitude and difference waves by group and condition. Difference waves at Fz are plotted for controls (green) and amusics (purple) in each condition (A-D). Purple and green colored regions show ± 1 SEM. Shaded gray regions in (A) and (C) show significant MMN at Fz across groups in response to harmonic and no-beating stimuli, while dashed gray rectangles in (B) and (D) show these same clusters for comparison for inharmonic and beating sounds, where no MMN was observed. The shaded orange region in (A) shows a significant P3a for harmonic stimuli, while the dashed orange rectangles in (B-D) show this same cluster for comparison for other sounds where no P3a was observed. (E) shows mean MMN amplitude, averaged within the relevant spatiotemporal cluster for each group and condition. Error bars show ± 1 SEM. For harmonic and no-beating stimuli, averages were computed in the cluster for that stimulus across both groups. As no significant clusters were identified for dissonant stimuli, inharmonic and beating stimuli borrowed the clusters for harmonic and no-beating stimuli, respectively. (F) is as (E), but for P3a amplitude, with other sounds borrowing the cluster for harmonic stimuli.

We also observed an interaction between group (amusic vs. control) and feature (beating vs. harmonicity) [$F(1,38) = 4.24, p = 0.046, \eta_p^2 = 0.10$]. The direction of the interaction (see Figure 7) suggests that MMNs in the control group were larger for harmonicity cues (HDEV and IDEV) than for beating cues (NDEV and BDEV), but MMNs in the amusic group were larger for beating than harmonicity. However, post-hoc between-group comparisons (independent t-tests) showed no significant differences between the amusic and control groups for harmonicity cues [$t(1,38) = -1.26, p = 0.22$] and only a marginally significant difference for beating [$t(1,38) = 1.78, p = 0.084$]. Within-group comparisons (paired t-tests) also showed no significant

differences between MMN amplitude to harmonicity and beating cues in either the amusic [$t(1,18) = -1.68, p = 0.11$] or control [$t(1,20) = 1.17, p = 0.25$] groups. The interaction was not significant if the 13 participants who completed Experiment 1 were removed [$F(1,25) = 3.16, p = 0.088, \eta_p^2 = 0.11$], nor when the analysis was restricted to include only consonant-deviant conditions, where significant MMNs were observed [$F(1,38) = 1.78, p = 0.19, \eta_p^2 = 0.06$].

No other significant main effects or interactions were identified in the ANOVA, regardless of whether the 13 participants who completed Experiment 1 were removed (all $p > 0.57$). The notable lack of main effect of group means that we found no evidence for reduced MMNs to consonance cues in amusic individuals overall. This neurophysiological outcome is in contrast to our behavioral results, where amusics were impaired overall for both cues, both in terms of threshold and sensitivity.

In order to confirm the presence of an MMN for each group in each condition, we compared averaged MMN strengths against zero using one-tailed t-tests, since the expected direction of the effect is negative. Only three MMNs were significant using this test: of controls to harmonic stimuli [$t(1,20) = -3.06, p = 0.0031$], of controls to no-beating stimuli [$t(1,20) = -1.86, p = 0.039$], and of amusics to no-beating stimuli: [$t(1,18) = -2.13, p = 0.024$]. These results, combined with the observed interaction effect, suggest a comparative advantage, in terms of early brain responses, for harmonicity cues for controls and beating cues for amusics.

To evaluate P3a amplitudes across group and condition, we ran a 3-way mixed ANOVA on ERP amplitudes in the P3a cluster identified for harmonic stimuli. No main effects or interactions were observed (all $p > 0.12$). Limiting the analysis to harmonic stimuli where a P3a response was observed, we observed no significant difference in P3a amplitude between amusic and control groups [$t(1,38) = 0.09, p = 0.93$]. To verify the presence of a P3a response for each

group in each condition, we compared averaged P3a strengths against zero using one-tailed t-tests. Only the P3a responses to harmonic stimuli were significant, for both controls [$t(1,20) = 2.16, p = 0.022$], and amusics [$t(1,18) = 1.77, p = 0.047$].

4. DISCUSSION

In our study, amusic participants showed an impaired ability to discriminate dissonant from consonant sounds, whether using beating or inharmonicity cues. Their MMN responses to the same stimuli were of similar amplitude to those of controls, although they showed a comparatively stronger MMN to beating cues than to inharmonicity cues, relative to controls. These findings have implications for understanding congenital amusia, as well as for the processing of consonance and dissonance overall.

4.1 Processing of harmonicity and beating in congenital amusia

We observed that amusic individuals have elevated thresholds (Experiment 1) and decreased behavioral sensitivity (Experiment 2) for detection of beating as well as of inharmonicity. Although amusics are behaviorally less sensitive to both beating and harmonicity cues, we found that their MMN responses to these cues were of comparable overall strength to control participants. This finding suggests that initial encoding of these dissonance cues may be intact in amusics despite their poor behavioral performance, pointing to higher-level mechanisms as the source for behavioral impairment. Our results thus agree with previous research suggesting that amusia-related impairments originate in a frontotemporal network connecting auditory areas to higher-order areas involved in working memory and conscious judgments (Albouy et al., 2013, 2015; Hyde et al., 2006, 2007, 2011; Leveque et al., 2016).

Our general finding of behavioral impairment for the perception of consonance in amusia agrees with previous studies (Cousineau et al., 2012; Marin et al., 2015) that used different methods and stimuli to address a similar question. In these studies, participants gave subjective pleasantness ratings in response to more ecologically valid musical stimuli. These stimuli have the advantage of being directly relevant to the real-world experience of music by people with amusia: we can directly conclude that they do not perceive consonance in the same way as control participants. In the present study, as well as in Experiments 3 and 4 of Cousineau et al. (2012), participants were instead tested with objective discrimination tasks, using more well controlled, but highly artificial, lab-synthesized stimuli. On the general question of impairment for consonance/dissonance sensitivity in amusia, the present study agrees with these past studies.

Contrary to the conclusions of Cousineau et al. (2012), however, our findings suggest that congenital amusia involves deficits in the perception of beating as well as the perception of harmonicity. The difference between the previous and present studies likely arises from the fact that the present study tested participants at low modulation depths, at and near threshold, as opposed to using high modulation depths that were easily detectable for both groups (as in Cousineau et al., 2012). Our finding of behavioral impairment for beating detection (and closely related AM detection) replicates a recent finding that amusic individuals exhibit deficits for AM as well as FM coding (Whiteford & Oxenham, 2017), also using threshold measurements.

Nevertheless, the interaction effect we observed in MMN amplitudes suggests that amusics' encoding of beating may be stronger relative to harmonicity, compared to controls. Although this interaction was not significant when analyses were limited to conditions with significantly negative MMNs, the tendency remains and is supported by the finding of a

significant MMN to harmonic sounds for controls, but not for amusics. This interaction may in theory reflect either impaired harmonic encoding or enhanced AM encoding in amusia. Given the behavioral results, both here and in a previous study (Whiteford & Oxenham, 2017), suggesting that amusia is related to impairments in both inharmonicity and beat detection, the theory of impaired harmonic perception in amusia seems more plausible. Nevertheless, despite these behavioral deficits, our EEG findings suggest that amusics are, at a minimum, not significantly impaired for early processing of AM in the form of beating cues. It might be argued that the increased representation of beating cues in the MMN of amusic individuals could be a primary, congenital difference in brain function of amusic individuals. But perhaps more likely, it may be a secondary result of impoverished spectral information, in which amusics develop relatively heightened representations of early information from sources that remain available. This phenomenon may also explain why amusic individuals remain able to use acoustic cues, such as loudness (Graves et al., 2019) or non-spectral cues for prosody (Pralus et al., 2019). Under this explanation, the increased emphasis on beating cues could be compared to some examples of enhanced visual perception for people with profound hearing loss (e.g. Bernstein et al., 2000; Bottari et al., 2011), or to enhanced auditory perception for people with low vision (e.g. Kolarik, Cirstea, Pardhan, & Moore, 2014). It must be noted, however, that the amusic individuals in our study did not show increased perceptual sensitivity to beating cues relative to controls, but merely a stronger relative importance of beating over harmonicity in early processing.

Recent evidence suggests that the unpleasantness of dissonant sounds is largely determined by culture, expertise, and context within a musical style, but that even when rated as pleasant, the tension of dissonant sounds can be perceived, allowing them to be discriminated

from consonant sounds (Lahdelma & Eerola, 2020; Popescu et al., 2019). Studies such as the present one, which rely on objective discriminability of consonant and dissonant sounds, do not inform us whether the perception of dissonance is subjectively unpleasant, but merely that it can be discriminated.

4.2 Beating and harmonicity cues in consonance and dissonance perception

Behaviorally, participants in our study detected dissonant deviants more easily in a stream of consonant standards than the reverse. This effect is in line with previous evidence for a behavioral advantage for detecting dissonance (Schellenberg & Trehub, 1994), an effect which has also been observed in infants (Schellenberg & Trehub, 1996). More generally, the advantage for detecting dissonance may be an example of perceptual asymmetry due to feature extraction. In both auditory (e.g. Cusack & Carlyon, 2003; Ruggles & Oxenham, 2014) and visual perception (e.g. Treisman & Gormican, 1988), observers are better able to detect the presence of an extracted feature than its absence. The behavioral advantage for dissonance detection suggests that dissonance cues (beating and inharmonicity) may be extracted as features by the auditory system, whereas their opposites (lack of beating and harmonicity) are only detectable as the absence of a feature.

More surprising than the behavioral effect is the MMN advantage we observed for consonant deviants over dissonant deviants, which runs in the opposite direction to behavioral results. It also runs opposite a recently reported effect of larger MMNs for dissonant deviants than for consonant deviants (Crespo-Bojorque et al., 2018). However, the stimuli in that study were slightly different from those in the present study: whereas the previous study used pairs of simultaneous tones forming consonant or dissonant harmonic intervals, the present study used single complex tones containing isolated low-level dissonance cues. The discrepancy between

the present study and Crespo-Bojorque et al. (2018) is perhaps due to the fact that all tones in that study were individually harmonic, although dissonant combinations produced inharmonicity and beating in their interactions. A recent study also found larger MMN amplitudes for harmonic sounds than for inharmonic sounds in amusics and controls (Quiroga-Martinez et al., 2022). In that study, deviants and standards were either both harmonic or both inharmonic, unlike the present study in which deviants differed from standards in terms of harmonicity. Nevertheless, Quiroga-Martinez et al. (2022) provide converging evidence for the asymmetry in processing of harmonic and inharmonic sounds. An earlier study also supports our finding of a clear MMN to harmonic deviants: Jones (2003) presented rare, deviant harmonic complexes in a stream of mostly inharmonic complexes, in a condition comparable to HDEV in the present study, and observed an MMN. However, as the opposite condition was not tested in that paper, it provides no specific agreement with the present study's finding for increased MMN to consonant deviants over dissonant deviants. In the present study, the MMN advantage for consonant deviants did not interact with group, suggesting that the asymmetry is present for both amusics and controls. In any case, the direction of the MMN advantage for consonant or dissonant stimuli has no direct bearing on the main findings of the present paper, as an MMN in either case indicates an early brain response indexing detection of the change.

The advantage we observed for behavioral detection of dissonance, confirming previous effects (Schellenberg & Trehub, 1994, 1996), seems overall easier to explain than the opposite MMN advantage for consonant deviants, for which previous evidence is mixed (Crespo-Bojorque et al., 2018; Jones, 2003; Quiroga-Martinez et al., 2022). It may be that early processing of consonance is not guided by the same incentives as later conscious processing of consonance, leading to opposite advantages for consonant and dissonant deviants at these two

stages of processing. In early stages, the MMN advantage for consonant deviants might be incentivized due to asymmetry in the entropy of harmonic and inharmonic signals, as suggested by Quiroga-Martinez et al. (2022). Within a predictive coding framework (Kanai et al., 2015; Rao & Ballard, 1999), increased MMNs for consonant deviants may reflect a higher precision and thus increased weight of the prediction error, due to the lower entropy of a harmonic deviant signal in comparison with the higher entropy of an inharmonic deviant signal. Thus, even if both kinds of deviants violate predictions of equal strength, consonant deviants may receive increased weight as prediction errors due to their decreased entropy. Because consonant sounds tend to have more salient pitch, a larger response to consonant deviants may also be related to the pitch onset response, an observed increase in activity in Heschl's gyrus at the onset of periodic sound, with no equivalent response at pitch offset (Gutschalk et al., 2004; Krumbholz et al., 2003). But if consonant deviants are more salient in early processing, how can we explain the behavioral advantage for dissonant deviants? There are several reasons dissonant deviants may be easier to detect when applied to real-world environments: dissonance may be more likely to indicate a threat, and dissonance may be overall rarer in the auditory world. There may also be an effect of long-term experience, whereby listeners have much more practice identifying dissonance against a consonant backdrop than the reverse.

4.3 Limitations in measuring the MMN

The present study focused on stimuli with low levels of dissonance by measuring thresholds in Experiment 1 and presenting stimuli for MMN measurements in Experiment 2 at constant differences near threshold. Although this approach allowed us to detect a behavioral impairment for beating perception in amusic individuals that was not detected by Cousineau et al. (2012), it also meant that our analyses of behavioral sensitivity in the MMN experiment were

conducted on data near chance performance, and that our EEG analyses were conducted near the threshold for emergence of MMN. The noise inherent in both of these measurements at such low levels may have hampered our statistical ability to detect differences between groups or conditions, as behavioral performance above chance and the presence of an MMN were not confirmed for every group in every condition. This was especially true for the amusic participants, whose behavioral sensitivity for oddball detection in Experiment 2 was only significantly above chance in one condition out of four (BDEV). Future studies might further explore this phenomenon by presenting stimuli with larger levels of dissonance in order to observe a more robust MMN response. It may also be considered a limitation of our study that not all participants completed both experiments, meaning that thresholds were known for some but not all participants in Experiment 2.

Another possible reason for small overall MMN amplitudes in the present study was the variability of the standard stimulus, which changed in F0 for each stimulus presentation. The variability of F0 was necessary in our study to ensure that responses could be interpreted as detection of dissonance, and not as perception of a change in pitch. However, most MMN studies are done using a constant, invariable standard, and the variability in our standard stimulus may have led to decreased MMN amplitude. In addition, stimuli in our study were largely dissonant overall, because three out of four tone types were inharmonic (inharmonic, no-beating, and beating). This was a necessary control measure to ensure that beating stimuli were detected using beats and not using the presence of inharmonic components. However, it may have reduced the salience of the consonance contrast along the beating axis, since both tone types on that axis are dissonant.

4.4 Conclusions

Our present findings revealed reduced behavioral sensitivity in amusic participants to two cues that have been implicated in consonance and dissonance perception: beating and inharmonicity. However, our EEG results also suggest that the MMN was not impaired overall in amusic relative to control participants, but inharmonicity and beating may be differently impacted in congenital amusia, with beating cues given relatively more weight than harmonicity cues in participants with amusia. Taken together, our findings are compatible with the view that inharmonicity and beating each contribute to dissonance perception, but that the relative weight of beating cues over inharmonicity cues may be increased for amusic individuals in comparison to control participants.

5. ACKNOWLEDGMENTS

This work was supported by a grant from the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience, by LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-0060), LabEx Cortex ("Construction, Function and Cognitive Function and Rehabilitation of the Cortex", ANR-11-LABX-0042) of Université de Lyon, within the program "Investissements d'avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR), by the AMBISENSE ("Access to Perceptual Ambiguity") grant from the French National Research Agency (ANR), and by NIH grant R01 DC005216.

6. REFERENCES

Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: evidence from event-related brain potentials. In *Journal of experimental psychology. Human perception and performance* (Vol. 27, Issue 5, pp. 1072–1089).

<https://doi.org/10.1037/0096-1523.27.5.1072>

Albouy, P., Caclin, A., Norman-Haignere, S. V., Lévêque, Y., Peretz, I., Tillmann, B., & Zatorre, R. J. (2019). Decoding task-related functional brain imaging data to identify developmental disorders: The case of congenital amusia. *Frontiers in Neuroscience, 13*(OCT), 1–13.

<https://doi.org/10.3389/fnins.2019.01165>

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*(5), 1639–1661.

<https://doi.org/10.1093/brain/awt082>

Albouy, P., Mattout, J., Sanchez, G., Tillmann, B., & Caclin, A. (2015). Altered retrieval of melodic information in congenital amusia : insights from dynamic causal modeling of MEG data. *Frontiers in Human Neuroscience, 9*, 1–13. <https://doi.org/10.3389/fnhum.2015.00020>

Albouy, P., Peretz, I., Bermudez, Pa., Zatorre, R. J., Tillmann, B., & Caclin, A. (2019).

Specialized neural dynamics for verbal and tonal memory : fMRI evidence in congenital amusia. *Human Brain Mapping, 1*–13. <https://doi.org/10.1002/hbm.24416>

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

<http://www.ncbi.nlm.nih.gov/pubmed/11844725>

Bernstein, L. E., Demorest, M. E., & Tucker, P. E. (2000). Speech perception without hearing.

Perception & Psychophysics, 62(2), 233–252.

<http://www.ncbi.nlm.nih.gov/pubmed/10723205>

Blair, R. C., & Karniski, W. (1993). An alternative method for significance testing of waveform difference potentials. *Psychophysiology, 30*(5), 518–524. <https://doi.org/10.1111/j.1469->

8986.1993.tb02075.x

Bottari, D., Caclin, A., Giard, M. H., & Pavani, F. (2011). Changes in early cortical visual processing predict enhanced reactivity in deaf individuals. *PLoS ONE*, *6*(9).

<https://doi.org/10.1371/journal.pone.0025607>

Brattico, E., Pallesen, K. J., Varyagina, O., Bailey, C., Anourova, I., Järvenpää, M., Eerola, T., & Tervaniemi, M. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *Journal of Cognitive Neuroscience*, *21*(11), 2230–2244.

<https://doi.org/10.1162/jocn.2008.21144>

Butler, B. E., & Trainor, L. J. (2012). Sequencing the cortical processing of pitch-evoking stimuli using EEG analysis and source estimation. *Frontiers in Psychology*, *3*(JUN), 1–13.

<https://doi.org/10.3389/fpsyg.2012.00180>

Butler, J. W., & Daston, P. G. (1968). Musical consonance as musical preference: a cross-cultural study. *The Journal of General Psychology*, *79*(1), 129–142.

<https://doi.org/10.1080/00221309.1968.9710460>

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. In *Statistical Power Analysis for the Behavioral Sciences*. Routledge Academic.

<https://doi.org/10.4324/9780203771587>

Cousineau, M., McDermott, J. H., & Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(48), 19858–19863. <https://doi.org/10.1073/pnas.1207989109>

Cousineau, M., Oxenham, A. J., & Peretz, I. (2015). Congenital amusia: a cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia*, *66*, 293–301. <https://doi.org/10.1002/ana.22528>. Toll-like

- Crespo-Bojorque, P., Monte-Ordoño, J., & Toro, J. M. (2018). Early neural responses underlie advantages for consonance over dissonance. *Neuropsychologia*, *117*(June), 188–198.
<https://doi.org/10.1016/j.neuropsychologia.2018.06.005>
- Cusack, R., & Carlyon, R. P. (2003). Perceptual asymmetries in audition. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(3), 713–725.
<https://doi.org/10.1037/0096-1523.29.3.713>
- De Baene, W., Vandierendonck, A., Leman, M., Widmann, A., & Tervaniemi, M. (2004). Roughness perception in sounds: Behavioral and ERP evidence. *Biological Psychology*, *67*(3), 319–330. <https://doi.org/10.1016/j.biopsycho.2004.01.003>
- 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.
- Ebeling, M. (2008). Neuronal periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion. *The Journal of the Acoustical Society of America*, *124*(4), 2320–2329. <https://doi.org/10.1121/1.2968688>
- Foxton, J. M., Nandy, R. K., & Griffiths, T. D. (2006). Rhythm deficits in “tone deafness”. *Brain and Cognition*, *62*(1), 24–29.
- Graves, J. E., Pralus, A., Fornoni, L., Oxenham, A. J., Caclin, A., & Tillmann, B. (2019). Short- and long-term memory for pitch and non-pitch contours : Insights from congenital amusia. *Brain and Cognition*, *136*, 103614. <https://doi.org/10.1016/j.bandc.2019.103614>
- Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2004). Temporal dynamics of pitch in human auditory cortex. *NeuroImage*, *22*(2), 755–766.
<https://doi.org/10.1016/j.neuroimage.2004.01.025>

Hutchinson, W., & Knopoff, L. (1978). The acoustic component of western consonance.

Interface, 7(1), 1–29. <https://doi.org/10.1080/09298217808570246>

Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007).

Cortical thickness in congenital amusia: when less is better than more. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 27(47), 13028–13032.

<https://doi.org/10.1523/JNEUROSCI.3039-07.2007>

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, 15(5), 356–360.

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.

<https://doi.org/10.1093/brain/awl204>

Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex*, 21(February), 292–299. <https://doi.org/10.1093/cercor/bhq094>

Jones, S. J. (2003). Sensitivity of human auditory evoked potentials to the harmonicity of complex tones: Evidence for dissociated cortical processes of spectral and periodicity analysis. *Experimental Brain Research*, 150(4), 506–514. <https://doi.org/10.1007/s00221-003-1482-9>

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

<https://doi.org/10.1111/j.1469-1809.1980.tb01571.x>

Kanai, R., Komura, Y., Shipp, S., & Friston, K. (2015). Cerebral hierarchies: Predictive

processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society B:*

- Biological Sciences*, 370(1668). <https://doi.org/10.1098/rstb.2014.0169>
- Kohlrausch, A., Fassel, R., & Dau, T. (2000). The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *The Journal of the Acoustical Society of America*, 108(2), 723–734. <https://doi.org/10.1121/1.429605>
- Kolarik, A. J., Cirstea, S., Pardhan, S., & Moore, B. C. J. (2014). A summary of research investigating echolocation abilities of blind and sighted humans. *Hearing Research*, 310(0), 60–68. <https://doi.org/10.1016/j.heares.2014.01.010>
- Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., & Lütkenhöner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, 13(7), 765–772. <https://doi.org/10.1093/cercor/13.7.765>
- Lahdelma, I., & Eerola, T. (2020). Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Scientific Reports*, 10(1), 1–11. <https://doi.org/10.1038/s41598-020-65615-8>
- Leveque, Y., Fauvel, B., Groussard, M., Caclin, A., Albouy, P., Platel, H., & Tillmann, B. (2016). Altered intrinsic connectivity of the auditory cortex in congenital amusia. *Journal of Neurophysiology*, 116(1), 88–97. <https://doi.org/10.1152/jn.00663.2015>
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2), 467–477. <https://doi.org/10.1121/1.1912375>
- Linnavalli, T., Ojala, J., Haveri, L., Putkinen, V., Kostilainen, K., Seppänen, S., & Tervaniemi, M. (2020). Musical expertise facilitates dissonance detection on behavioral, not on early sensory level. *Music Perception*, 38(1), 78–98. <https://doi.org/10.1525/mp.2020.38.1.78>
- Lundin, R. W. (1947). Toward a cultural theory of consonance. *Journal of Psychology: Interdisciplinary and Applied*, 23(1), 45–49.

<https://doi.org/10.1080/00223980.1947.9917318>

Marin, M. M., Gingras, B., & Stewart, L. (2012). Perception of musical timbre in congenital amusia: Categorization, discrimination and short-term memory. *Neuropsychologia*, *50*(3), 367–378. <https://doi.org/10.1016/j.neuropsychologia.2011.12.006>

Marin, M. M., Thompson, W. F., Gingras, B., & Stewart, L. (2015). Affective evaluation of simultaneous tone combinations in congenital amusia. *Neuropsychologia*, *78*, 207–220. <https://doi.org/10.1016/j.neuropsychologia.2015.10.004>

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>

Marmel, F., Perrin, F., & Tillmann, B. (2011). Tonal expectations influence early pitch processing. *Journal of Cognitive Neuroscience*, *23*, 3095–3104. <https://doi.org/10.1162/jocn.2011.21632>

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, *20*(11), 1035–1041. <https://doi.org/10.1016/j.cub.2010.04.019> Individual

McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, *535*(7613), 547–550. <https://doi.org/10.1038/nature18635>

Moreau, P., Jolicœur, P., & Peretz, I. (2009). Automatic brain responses to pitch changes in congenital amusia. *Annals of the New York Academy of Sciences*, *1169*, 191–194. <https://doi.org/10.1111/j.1749-6632.2009.04775.x>

Moreau, P., Jolicœur, P., & Peretz, I. (2013). Pitch discrimination without awareness in

- congenital amusia: Evidence from event-related potentials. *Brain and Cognition*, 81(3), 337–344. <https://doi.org/10.1016/j.bandc.2013.01.004>
- Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313–329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- 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. <https://doi.org/10.1016/j.clinph.2007.04.026>
- Noguchi, K., Gel, Y. R., Brunner, E., & Konietzschke, F. (2012). nparLD : An R Software Package for the Nonparametric Analysis of Longitudinal Data in Factorial Experiments . *Journal of Statistical Software*, 50(12). <https://doi.org/10.18637/jss.v050.i12>
- Norman-Haignere, S. V., Albouy, P., Caclin, A., Mcdermott, J. H., Kanwisher, N. G., & Tillmann, B. (2016). Pitch-responsive cortical regions in congenital amusia. *The Journal of Neuroscience*, 36(10), 2986–2994. <https://doi.org/10.1523/JNEUROSCI.2705-15.2016>
- Omigie, D., Pearce, M. T., Williamson, V. J., & Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. *Neuropsychologia*, 51(9), 1749–1762. <https://doi.org/10.1016/j.neuropsychologia.2013.05.010>
- Peretz, I. (2016). Neurobiology of congenital amusia. *Trends in Cognitive Sciences*, 20(11), 857–867. <https://doi.org/10.1016/j.tics.2016.09.002>
- Peretz, I., Brattico, E., Järvenp, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain*, 132(5), 1277–1286. <https://doi.org/10.1093/brain/awp055>
- Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, 999, 58–75.

- Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal of Human Genetics*, 25(5), 625–630. <https://doi.org/10.1038/ejhg.2017.15>
- Plomp, R., & Levelt, W. J. (1965). Tonal consonance and critical bandwidth. *The Journal of the Acoustical Society of America*, 38(4), 548–560.
<http://www.ncbi.nlm.nih.gov/pubmed/5831012>
- Popescu, T., Neuser, M. P., Neuwirth, M., Bravo, F., Mende, W., Boneh, O., Moss, F. C., & Rohrmeier, M. (2019). The pleasantness of sensory dissonance is mediated by musical style and expertise. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-018-35873-8>
- Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., & Caclin, A. (2019). Emotional prosody in congenital amusia: Impaired and spared processes. *Neuropsychologia*, 134(July 2018), 107234.
<https://doi.org/10.1016/j.neuropsychologia.2019.107234>
- Pralus, A., Gomot, M., Graves, J., Cholvy, F., Fornoni, L., Tillmann, B., & Caclin, A. (2020). Pre-attentive processing of neutral and emotional sounds in congenital amusia. *BioRxiv*, 1–43.
- Quiroga-Martinez, D. R., Basiński, K., Nasielski, J., Tillmann, B., Brattico, E., Cholvy, F., Fornoni, L., Vuust, P., & Caclin, A. (2022). Enhanced mismatch negativity in harmonic compared with inharmonic sounds. *European Journal of Neuroscience*, 56(5), 4583–4599.
<https://doi.org/10.1111/ejn.15769>
- Quiroga-Martinez, D. R., Tillmann, B., Brattico, E., Cholvy, F., Fornoni, L., Vuust, P., & Caclin, A. (2021). Listeners with congenital amusia are sensitive to context uncertainty in melodic sequences. *Neuropsychologia*, 158(June), 107911.
<https://doi.org/10.1016/j.neuropsychologia.2021.107911>

- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. <https://doi.org/10.1038/4580>
- Ruggles, D. R., & Oxenham, A. J. (2014). Perceptual asymmetry induced by the auditory continuity illusion. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 908–914. <https://doi.org/10.1037/a0035411>
- Schellenberg, E. G., & Trehub, S. E. (1994). Frequency ratios and the discrimination of pure tone sequences. *Perception & Psychophysics*, *56*(4), 472–478. <https://doi.org/10.3758/BF03206738>
- Schellenberg, E. G., & Trehub, S. E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science*, *7*(5), 272–277. <https://doi.org/10.1111/j.1467-9280.1996.tb00373.x>
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc.*, *31*(1), 137–149. <http://www.ncbi.nlm.nih.gov/pubmed/10495845>
- Stumpf, C. (1883). *Tonpsychologie: Volume I*. Cambridge University Press.
- Terhardt, E. (1974). Pitch, consonance, and harmony. *The Journal of the Acoustical Society of America*, *55*(5), 1061–1069. <https://doi.org/10.1121/1.1914648>
- Tervaniemi, M., Schröger, E., Saher, M., & Näätänen, R. (2000). Effects of spectral complexity and sound duration on automatic complex-sound pitch processing in humans - A mismatch negativity study. *Neuroscience Letters*, *290*(1), 66–70. [https://doi.org/10.1016/S0304-3940\(00\)01290-8](https://doi.org/10.1016/S0304-3940(00)01290-8)
- Tillmann, B., Albouy, P., & Caclin, A. (2015). Congenital amusias. In *Handbook of Clinical*

Neurology (1st ed., Vol. 129). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-62630-1.00033-0>

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

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*(1), 15–48. <https://doi.org/10.1037//0033-295x.95.1.15>

von Helmholtz, H. (1912). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Cambridge University Press.

Whiteford, K. L., & Oxenham, A. J. (2017). Auditory deficits in amusia extend beyond poor pitch perception. *Neuropsychologia, 99*(October 2016), 213–224. <https://doi.org/10.1016/j.neuropsychologia.2017.03.018>

Figure S1. Thresholds for pure-tone AM detection. Thresholds are shown with average and SEM (left), and as individual data compared against single-sideband beating detection thresholds (middle) and against inharmonic shift thresholds (right). Significant correlations were observed between AM and Beating thresholds [$r = 0.73$, $p < 0.001$], and between AM and Harmonicity thresholds [$r = 0.56$, $p = 0.005$].

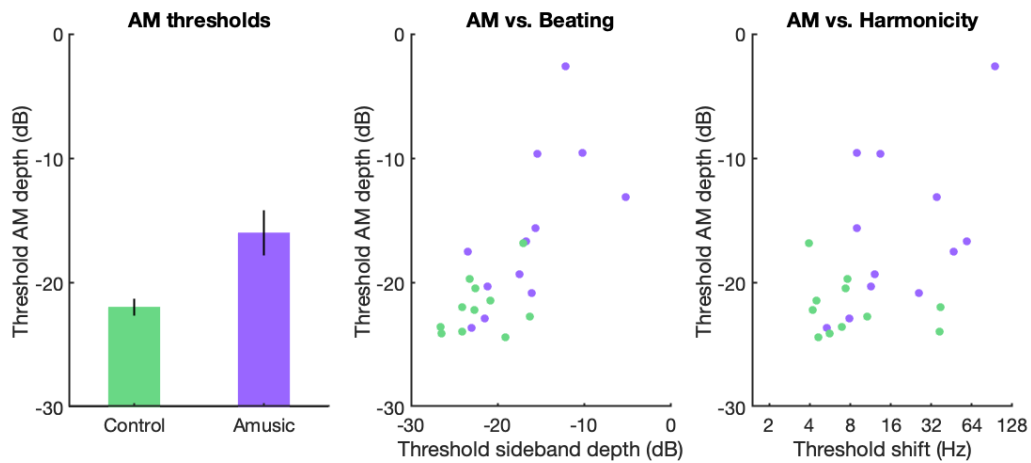


Figure S2. **Topography of MMN and P3a responses to dissonant stimuli.** As figure 5, but for dissonant stimuli (inharmonic and beating), where no significant MMNs or P3as were observed (as opposed to consonant stimuli shown in figure 5, where significant differences were found).

