Published in final edited form as:

Neuropsychologia. 2017 May ; 99: 213–224. doi:10.1016/j.neuropsychologia.2017.03.018.

Auditory Deficits in Amusia Extend Beyond Poor Pitch Perception

Kelly L. Whiteford* and Andrew J. Oxenham

Department of Psychology, University of Minnesota, 75 East River Parkway, Minneapolis, MN 55455, USA

Abstract

Congenital amusia is a music perception disorder believed to reflect a deficit in fine-grained pitch perception and/or short-term or working memory for pitch. Because most measures of pitch perception include memory and segmentation components, it has been difficult to determine the true extent of pitch processing deficits in amusia. It is also unclear whether pitch deficits persist at frequencies beyond the range of musical pitch. To address these questions, experiments were conducted with amusics and matched controls, manipulating both the stimuli and the task demands. First, we assessed pitch discrimination at low (500 Hz and 2000 Hz) and high (8000 Hz) frequencies using a three-interval forced-choice task. Amusics exhibited deficits even at the highest frequency, which lies beyond the existence region of musical pitch. Next, we assessed the extent to which frequency coding deficits persist in one- and two-interval frequency-modulation (FM) and amplitude-modulation (AM) detection tasks at 500 Hz at slow ($f_m = 4$ Hz) and fast ($f_m = 20$ Hz) modulation rates. Amusics still exhibited deficits in one-interval FM detection tasks that should not involve memory or segmentation. Surprisingly, amusics were also impaired on AM detection, which should not involve pitch processing. Finally, direct comparisons between the detection of continuous and discrete FM demonstrated that amusics suffer deficits both in coding and segmenting pitch information. Our results reveal auditory deficits in amusia extending beyond pitch perception that are subtle when controlling for memory and segmentation, and are likely exacerbated in more complex contexts such as musical listening.

Keywords

amusia; pitch discrimination; memory; frequency modulation; amplitude modulation; auditory object perception

1. Introduction

Pitch, the psychological attribute that allows us to order sounds on a musical scale from low to high (ANSI, 2013), plays a fundamental role in our auditory worlds, helping us to

*Corresponding author. whit1945@umn.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

segregate sounds, recognize voices, and enjoy our favorite music. Pitch is related to the repetition rate, or periodicity, of an acoustic waveform. For pure tones, or sinusoids that just have one frequency, the strength and accuracy of the tone's pitch depends on attributes such as its duration and frequency (Micheyl et al., 2012; Moore, 1973; Moore and Ernst, 2012).

Even with the same stimulus parameters, large differences in pitch discrimination abilities have been observed between individuals, due to factors such as experience (e.g., Carcagno and Plack, 2011; Micheyl et al., 2006), age (e.g., Moore and Peters, 1992), hearing loss (Arehart, 1994; Moore and Moore, 2003; Moore and Peters, 1992), and neurogenetic disorders (Peretz et al., 2007). One such disorder known to be related to deficits in pitch perception is congenital amusia (e.g., Vuvan et al., 2015). Amusia, commonly referred to as tone deafness, is a deficit in music perception that is believed to be independent of hearing loss, musical training, or intelligence (Ayotte et al., 2002; Peretz, 2001). Amusics typically report that they are "musically impaired," report difficulty recognizing familiar tunes when presented without lyrics, and cannot perceive when they or others are singing out of tune (Peretz et al., 2008, 2003). Amusia may be congenital (i.e., developed at birth or in early childhood) or can be acquired through brain injury to areas important for music perception (e.g., bilateral damage to superior temporal lobes) (Peretz, 2001; Peretz et al., 1994). Music perception deficits in amusia appear to be linked to an underlying impairment in fine-grained pitch perception (e.g., Foxton et al., 2004; Peretz et al., 2002; Vuvan et al., 2015), either due to deficits in pitch processing (e.g., Cousineau et al., 2015) or deficits in pitch memory (Albouy et al., 2015, 2013a; Gosselin et al., 2009; Tillmann et al., 2016, 2009; Williamson et al., 2010; Williamson and Stewart, 2010).

The exact nature of the pitch-processing deficits remains unclear. Cousineau et al. (2015) found that amusics exhibited a deficit in pitch discrimination for complex tones that contained low-numbered harmonics (e.g., 1–6), which are thought to be resolved in the auditory periphery and coded via their spectro-temporal fine structure. However, no deficit was observed in pitch discrimination for complex tones that contained only high-numbered harmonics (> 10), which are thought to be unresolved and coded via their temporal envelope (Bernstein and Oxenham, 2003; Houtsma and Smurzynski, 1990). Because amusics exhibited normal abilities to detect interaural time differences (ITD) via temporal fine structure, as well as normal auditory filter shapes, Cousineau et al. (2015) concluded that the amusics' deficit in coding the spectro-temporal fine structure of resolved harmonics was not peripheral in nature.

It is not known why pitch deficits related to amusia were observed with low-numbered resolved harmonics but not high-numbered unresolved harmonics. One possibility is that only pitch coding via temporal fine structure is affected, so that the (generally weaker) pitch elicited by temporal envelope cues is not affected. In other words, if different mechanisms are involved in coding resolved and unresolved harmonics (Carlyon and Shackleton, 1994; but see Micheyl and Oxenham, 2004), then amusia may selectively affect the coding of resolved harmonics, which provides the dominant pitch percept in everyday life (Plack and Oxenham, 2005). Another possibility is that amusic deficits are limited to fine-grained pitch differences (e.g., Foxton et al., 2004; Peretz et al., 2002; Vuvan et al., 2015), and because pitch discrimination of complexes with only unresolved harmonics is relatively coarse,

amusia deficits are no longer measurable in such conditions. Experiment 1 of our study addresses this question by measuring pitch discrimination for single pure tones over a range of frequencies, including a very high frequency (8 kHz), where discrimination is generally much poorer. If amusia deficits are linked to the poorer coding of individual harmonics and tones, then performance should remain poorer than normal even at high frequencies. On the other hand, if amusia deficits are limited solely to conditions where normal pitch discrimination is very fine, then the deficit should be reduced at 8 kHz, where normal pitch discrimination is degraded relative to discrimination at lower frequencies.

Another open question relating to amusia is the extent to which it affects perceptual (e.g., pitch processing) versus cognitive (e.g., short-term or working memory) processing. Attempts to find an anatomical or physiological marker for amusia have largely focused on cortical differences. One noted difference involves grey and white matter abnormalities in the right superior temporal gyrus (rSTG) and the right inferior frontal gyrus (rIFG) (Hyde et al., 2007, 2006), with amusics exhibiting increased grey matter in rSTG and rIFG relative to controls. This finding was somewhat counterintuitive, given that professional musicians have also been reported to have more grey matter relative to non-musicians in a subsection of the auditory cortex (e.g., anteromedial Heschl's gyrus; Schneider et al., 2002). Hyde et al. (2007) suggested that increases in grey matter in amusia may arise due to abnormal cortical migration during development. In line with the cortical migration theory, amusics have been found to exhibit impaired connectivity between the rIFG and the rSTG, confirmed via functional connectivity measures in fMRI (Hyde et al., 2011), resting-state fMRI (Lévesque et al., 2016), and Diffusion Tensor Imaging (Loui et al., 2009, although see Chen et al., 2015). More recently, abnormalities in connectivity patterns have been identified both within and between the auditory cortices using magneto-encephalography (MEG) (Albouy et al., 2013a, 2015b). These abnormalities within the frontotemporal network coincide with abnormal backward connectivity during pitch encoding (Albouy et al., 2013a) and abnormal forward connectivity during pitch retrieval (Albouy et al., 2015). Taken together, these results support the hypothesis that amusia is related to dysfunctions in the frontotemporal network, rather than abnormalities in early cortical or pre-cortical processing. Furthermore, recent fMRI results show no difference in the proportion, location, and selectivity of pitch-responsive voxels of the auditory cortex in amusics compared to controls (Norman-Haignere et al., 2016). Thus, current imaging findings might implicate deficits in memory and/or segmentation, rather than the initial processing of pitch. Although numerous studies provide evidence for impaired retention for short-term pitch information in amusia, whether or not amusics have a specific deficit in working memory for pitch (i.e., a difficulty in "online" comparison of pitch information over time) is an important question that has received little attention (Tillmann et al., 2016). Experiments 2 and 3 in this study provide behavioral tests to distinguish between these possibilities by comparing frequency-modulation (FM) detection in a single-interval task with more traditional frequency-discrimination paradigms involving one or more comparison intervals.

2. Experiment 1: Pure-tone frequency discrimination

The purpose of experiment 1 was to determine whether poor pitch perception in amusia extends to high frequencies. On one hand, amusia typically results in a poor ability to

discriminate small frequency differences (for a review, see Vuvan et al., 2015), so one might expect this deficit to be present at all frequencies. On the other hand, there are at least two considerations to suggest that the deficit may be limited to lower frequencies. The first consideration is that musical pitch perception (i.e., pitch sufficiently salient to convey musical intervals and melodies) only typically extends up to about 4–5 kHz (Attneave and Olson, 1971; Oxenham et al., 2011). If amusia is a deficit specifically related to musical pitch, then the deficit may not extend to high frequencies. The second consideration is that pitch deficits in amusia are limited to discrimination of small frequency differences of a semitone or less (~6%; e.g., Foxton et al., 2004; Hyde & Peretz, 2004). It may be that the advantage of control participants over amusics only holds in conditions where very fine discrimination is normally possible. Because the ability of normal participants to discriminate the pitch of pure tones worsens dramatically at high frequencies (e.g., Micheyl et al., 2012; Moore, 1973; Moore & Ernst, 2012), it may be that amusics perform more poorly at low frequencies but more similarly at high frequencies, as was found with spectrally resolved and unresolved components within a complex tone (Cousineau et al., 2015). We addressed this question by comparing the ability of amusic and normal control participants to discriminate changes in frequency of pure tones with frequencies of 500, 2000, and 8000 Hz.

2.1. Methods

2.1.1. Participants—Twelve amusics and 12 matched controls participated in the study. Participants were matched in age, years of musical experience, and years of education (see Table 1). In addition, IQ was measured via the Vocabulary and Matrix Reasoning subtests of the Wechsler Abbreviated Scale of Intelligence — Second Edition (WASI-II; Wechsler, 2011). The scores from the two subtests were combined to calculate the full scale IQ scores, which did not differ significantly between the groups (see Table 1). Because the WASI-II includes a measure of verbal intelligence, all participants were native speakers of American English. Amusia was determined based on a global score on the MBEA (Peretz et al., 2003) that was at least two standard deviations below the mean of the general population. All except one amusic participant also performed below the same cutoff for the pitch-based subtasks on the MBEA. None of the participants reported a history of neurological conditions. All participants provided written informed consent and were compensated with course credit or hourly payment for their time. The experimental protocols were approved by the Institutional Review Board of the University of Minnesota.

Audiometric thresholds were measured at octave frequencies from 250 to 8000 Hz. All except one participant had normal hearing up through 8 kHz, defined as audiometric thresholds no greater than 20 dB hearing level (HL). The one participant who did not meet this criterion was in the control group and had normal hearing up through 4 kHz, and a mild loss at 8 kHz (right ear: 25 dB HL; left ear: 35 dB HL). The average 8-kHz thresholds (across both ears) were not significantly different between the two groups ($t(22) = -.881$, $p = .388$, two-tailed).

2.1.2. Stimuli and procedures—The auditory tasks were administered in a double-walled sound-attenuating booth via Sennheiser HD650 headphones. All stimuli were

generated in MATLAB via a 24-bit L22 soundcard (LynxStudio, Costa Mesa, CA) with a sampling rate of 48,000 Hz and were presented at 60 dB sound pressure level (SPL).

Pure-tone frequency difference limens (FDLs) were measured at 500, 2000, and 8000 Hz. The tones were 200 ms in duration, including 50-ms raised-cosine onset and offset ramps. FDLs were obtained using a three-interval, three-alternative forced-choice (AFC) task with a two-down, one-up adaptive procedure that tracks the 70.7% correct point on the psychometric function (Levitt, 1971). On each trial, three tones were presented sequentially, separated by inter-stimulus intervals of 500 ms. Two of the tones were reference tones, identical in frequency. The other tone was the target tone, which was always higher in frequency than the reference tones. The frequencies of the reference and target tones were geometrically centered around the nominal test frequency of 500, 2000, or 8000 Hz. The presentation order of the target tone relative to the reference tones was selected randomly on each trial with uniform distribution. Subjects were instructed to pick the tone that was different by selecting one of three virtual buttons on the computer screen and to look at the computer screen to monitor the feedback, which was presented after each trial with the word “Correct” or “Incorrect”. Throughout the task, the three buttons were labeled “1”, “2”, and “3”. Each button was illuminated red during the presentation of the corresponding tone (i.e., button 1 was illuminated red during the presentation of the first tone, etc.). The starting value of the frequency difference (f) between the target and the reference tones was 20% (i.e., slightly greater than three semitones). The value of f varied by a factor of 2 for the first two reversal points, by a factor of 1.41 for the following two reversal points, and by a factor of 1.12 for the final six reversal points. The threshold for each run was calculated as the geometric mean of the f values at the last 6 reversal points. Each participant completed 2 runs at each frequency condition, for a total of 6 adaptive runs. The testing order of the frequency conditions was randomized across subjects and across repetitions, such that all three frequencies were tested before any was repeated. For the full experimental protocol involving all three experiments, see section S1 of the supplemental methods.

2.2. Results

Pure-tone FDLs are shown for amusics and controls in Fig. 1. Individual differences were large, and there was considerable overlap in FDLs between the two groups (Fig. 1A). However, on average, amusics had thresholds around 1 semitone (i.e., 6%) or greater, which was considerably higher than the mean thresholds observed in the control group (Fig. 1B). Interestingly, this trend was observed at all three frequencies tested, including 8000 Hz — a frequency that is typically too high to form recognizable melodies. A mixed-design ANOVA was performed, with log-transformed values of f (%) as the dependent variable, subject group as the between-subjects factor, and frequency as the within-subjects factor. Log-transforming FDLs is common practice to avoid compression close to zero and better approximate normality (e.g., Micheyl et al., 2012). Greenhouse-Geisser corrections were applied to correct for lack of sphericity, where appropriate (Frequency: $\varepsilon=.791$), and the corrected degrees of freedom are reported. The ANOVA revealed a main effect of group [$F(1,22) = 13.2, p = .001, \eta_p^2 = .375$], a main effect of frequency [$F(1.58,34.8) = 26.5, p < .0001, \eta_p^2 = .546$], and no interaction [$F(1.58,34.8) = .282, p = .705, \eta_p^2 = .013$] when using a criterion of $\alpha = .05$. Bonferroni corrected t -tests ($\alpha = .017$) indicated that thresholds at

8000 Hz were significantly higher (worse) than thresholds at 2000 Hz [$t(22) = 6.09, p < .0001$] and 500 Hz [$t(22) = 5.15, p < .0001$]. There was a non-significant trend for larger (worse) thresholds at 500 Hz compared to 2000 Hz [$t(22) = 1.80, p = .085$].

One amusic participant had unusually poor pitch perception at all frequencies, with FDLs as large as an octave (one octave = 100% FDL) or more. When the value of f was one octave, the two frequencies present in a trial were half an octave below and above the nominal test frequency, i.e., about 5.65 and 11.3 kHz, respectively, for a test frequency of 8 kHz. This participant reported having trouble remembering the first two tones by the time the third tone in the sequence was presented. Removal of this participant with unusually high FDLs, along with his/her matched control, did not change the main effects or lack of interaction reported above.

2.3. Discussion

The results suggest that poor frequency discrimination in amusics extends to 8 kHz, and thus beyond the traditional existence region of musical pitch. Pure-tone frequencies of 8 kHz are generally too high to form recognizable melodies or musical intervals (Attneave and Olson, 1971), and are much higher than even the highest musical note produced by a grand piano or the highest orchestral instrument (i.e., the piccolo). Therefore, our results suggest that amusia is not a deficit selective to the musical attributes of pitch.

Poor frequency discrimination at 8 kHz in amusics relative to controls corroborates previous findings of normal ITD discrimination in amusics, which suggest that amusia does not exclusively affect any putative temporal coding of pitch information in the auditory nerve (Cousineau et al., 2015). Even if low-frequency tones are coded via phase-locked information, it is generally believed that this timing information is no longer available at high frequencies. Because no direct measurements of auditory-nerve phase locking in humans exist, the cut-off frequency is unknown, but different estimates have ranged from around 1.5 kHz, based on the limits of sensitivity to ITDs (Joris and Verschooten, 2013), up to 4–5 kHz (Moore, 1973; Sek and Moore, 1995) or even higher (Moore and Ernst, 2012). However, there is general consensus that phase-locked information is highly unlikely to be available at 8 kHz. Thus, the fact that amusics show the same pattern of results as the normal control subjects suggests that their deficit is not limited to tones coded via peripheral temporal fine structure information, but rather may reflect a problem with central coding of the spectro-temporal fine structure associated with pure tones and spectrally resolved harmonics.

3. Experiment 2: Assessing the role of memory load

Performing frequency discrimination requires not only accurate coding of the frequency of each tone, but also memory storage and retrieval to allow successive tones to be compared (Durlach and Braida, 1969; Jesteadt and Sims, 1975; Zhang et al., 2016). It may be that poor FDLs in amusia are due not only to impaired frequency coding but also to deficits in short-term pitch memory (e.g., Albouy et al., 2013a; Gosselin et al., 2009; Williamson and Stewart, 2010) and/or manipulations of pitch information in working memory (Tillmann et al., 2016). There is conflicting evidence on whether amusia is related to problems with

short-term memory for pitch. Several studies note that short-term memory deficits in amusia persist even for stimuli that are above threshold (Albouy et al., 2013a, 2013b; Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010). However, presenting the same stimuli for both amusics and controls means that pitch changes will be closer to amusics' thresholds relative to controls (Jiang et al., 2013). Jiang et al. note it is difficult to interpret findings from memory studies that have not controlled for pitch sensitivity because pitch sensitivity and memory are confounded. Jiang et al. found that once controlling for each individual's perceptual sensitivity for pitch, there was no impairment in pitch memory for amusics relative to controls. In contrast, when perceptual sensitivity to pitch changes was not equated, Jiang et al. replicated findings from the previous studies (Albouy et al., 2013a; Gosselin et al., 2009; Williamson et al., 2010; Williamson and Stewart, 2010). They therefore concluded that amusia could not be attributed to deficits in short-term memory. However, even the paradigm of Jiang et al. (2013) relied to some extent on memory processes, as it still required a comparison of successive sounds across time. Auditory discrimination tasks may themselves utilize working memory if online comparisons are required to complete the task (Zhang et al., 2016). Therefore, Jiang et al. (2013) could have over-corrected for pitch sensitivity if their pitch change detection task required even minimal memory demands, as suggested by Tillmann et al. (2016).

To further alleviate memory load, we measured participants' sensitivity to frequency modulation (FM), using a one-interval yes/no task. After each tone, participants indicated "yes" if the tone was "changing" in pitch and "no" if the tone was not changing. Methods of signal detection theory were used to estimate sensitivity to FM in each group. An analogous task was used to measure amplitude modulation (AM) detection — a task that does not rely on pitch perception but requires coding of temporal envelope cues. As a comparison, FM and AM detection were also measured using a standard two-alternative forced choice (2AFC) procedure. In this task, participants were presented with two tones and were instructed to pick the tone that was changing. The 2AFC task presents the listener with more information than a single-interval task, but accessing the additional information again requires that the information from both intervals be stored and retrieved. If amusia reflects only memory-related deficits, then no deficits should be observed for tasks involving the one-interval detection of either FM or AM. On the other hand, if amusia involves a specific deficit in frequency coding, then a deficit should be observed for FM detection, but not for AM detection.

3.1. Methods

3.1.1. Participants—The same participants from experiment 1 also completed experiment 2.

3.1.2. Stimuli and procedure—Difference limens for AM (AMDLs) and FM (FMDLs) were measured for a 500-Hz pure-tone carrier at slow ($f_m = 4$ Hz) and fast ($f_m = 20$ Hz) modulation rates. Both the target and the unmodulated reference tone were 1.5 s in duration, including 50-ms raised-cosine onset and offset ramps. The level was roved between intervals in the range from 57 and 63 dB SPL, with uniform distribution, to avoid cues related to overall level or loudness. The presentation order of the conditions, modulation type (AM vs.

FM) and modulation rate ($f_m = 4$ Hz vs. $f_m = 20$ Hz), was counterbalanced across participants using a Latin Square design. Participants were yoked so that matched control-amusic pairs underwent the same counterbalanced order.

Before each adaptive run for the two-interval task, participants listened to an example of the corresponding target and reference tones. In order to ensure that the participants understood the type of modulation they would need to detect in the next experimental block of trials, they were required to correctly identify the example target by clicking a button labeled “Yes” after the modulated tone was presented. They were also required to correctly reject the reference tone by clicking a button labeled “No” after the pure tone was presented. Participants completed the example trials in the presence of the experimenter and were allowed to repeat the example trials as many times as necessary. The target for the AM example trials was presented at a modulation depth of 50% ($m = 0.5$), while the target for the FM example trials had a peak-to-peak frequency change of 12% (i.e., about 2 semitones). Most of the participants, including the amusics, did not need to repeat the example trials.

Initially, the two-interval, 2AFC task was used to estimate each individual’s sensitivity to AM and FM. Thresholds were measured using a three-down, one-up adaptive procedure that tracks the 79.4% correct point on the psychometric function (Levitt, 1971). On each trial, two tones were presented sequentially, separated by an interstimulus interval of 500 ms. One of the tones was the modulated target while the other was the reference tone with no modulation. Whether the target occurred in the first or second interval was selected randomly on each trial with equal *a priori* probability. The procedures for stimulus presentation and feedback were the same as for experiment 1. For the AM conditions, the starting value of the modulation depth was 50%. Modulation depth varied by a factor of 2 for the first two reversals, by a factor of 1.26 for the next two reversals, and by a factor of 1.12 for the final six reversals. For the FM conditions, the starting value of the frequency excursion from the carrier was 31.6 Hz. With a 500-Hz carrier, this corresponds to a peak-to-peak frequency change of approximately 12.6%, or 2 semitones. The frequency excursion varied by a factor of 2 for the first two reversals, by a factor of 1.41 for the next two reversals, and by a factor of 1.12 for the final six reversals. In all cases, the threshold for a given run was defined as the geometric mean of the tracking variable at the last 6 reversal points.

The thresholds from the 2AFC procedure were used to set the modulation depths for the one-interval AM task and the peak-to-peak frequency changes for the one-interval FM task. During the one-interval task, only one tone was presented in each trial. The task was to determine whether or not the tone was modulated by pressing the corresponding button on the screen. Subjects were instructed to press the “yes” button if they believed that the tone was modulated, and the “no” button if they did not believe that the tone was modulated. As before, example trials were presented before each modulation condition to ensure that the participants were familiar with the specific target they were to detect for the following block of trials. Because the target modulation was always presented near threshold in the experimental trials, the modulation depth and peak-to-peak frequency excursion of the target was half of that from the previous practice trials. After completing the corresponding

example trials, subjects were instructed to “listen carefully, as the modulation will be very subtle.” In the same manner as the two-interval task, modulation condition (AM vs. FM and modulation rate) was counterbalanced across participants using a Latin Square design.

Both the AM and FM one-interval tasks each involved a total of 175 trials per subject. In 20% of these trials the modulation depth was set at the individual’s threshold (i.e., the 79.4% correct point, based on the 2-interval adaptive procedure). In another 20% of the trials the modulation depth was above threshold by a factor of 1.59; in another 20% of trials the modulation depth was below the measured threshold by a factor of 1.58; and in another 20% of trials the modulation depth was set to be below threshold by a factor of 2.51. The remaining 20% of the trials contained unmodulated pure tones to provide an estimate of the false-alarm rate (catch trials). The presentation order of the tones with their modulation depths was random within each block of 175 trials.

3.2. Results

3.2.1. Comparing sensitivity to one- and two-interval FM and AM detection—

Analyses were conducted to compare one- and two-interval FM detection as well as one- and two-interval AM detection. Signal detection theory was used to calculate the sensitivity index, d' , for each observer on the one-interval FM and AM tasks. Assuming equal variance of the distributions underlying the response to the modulated and unmodulated stimuli, d' provides a criterion-free measure of sensitivity (Macmillan and Creelman, 2005). Proportions of 0 or 1 (i.e., all or no correctly identified modulation trials or no-modulation trials) were converted to $1/(2N)$ and $1-1/(2N)$, where N is the total number of trials on which the proportion is based (35 in this case), as suggested by Macmillan and Creelman (2005, pg. 8).

Four d' values were calculated for each participant, one for each of the four modulation indices (see Fig. 2 for an example), and were plotted as a function of the modulation index ($20\log(m)$ and $2\ f(\%)$) for AM and FM, respectively). The modulation index that corresponded to a d' of 1.14 (i.e., the sensitivity equivalent to the 79% correct point for a two-interval AFC task as used in the first-half of experiment 2) (Macmillan & Creelman, 2005, pg. 429) was estimated for each subject using linear regression. The fits for the linear regression were adequate; the mean and median R^2 for each modulation rate was always greater than 0.9.

Results from both groups in the two- and one-interval FM and AM tasks are presented in Fig. 3 (see Fig. S1 for individual data and text in supplementary section S2). Mixed-model ANOVAs were conducted on the log-transformed thresholds for both tasks, with modulation rate (4 and 20 Hz) and number of intervals (one vs. two) as the within-subjects factor and group (amusic and control) as the between-subjects factor. Any p values less than .05 were considered significant. The ANOVA for FM revealed a main effect of number of intervals [$F(1,22) = 9.34, p = .006, \eta_p^2 = .298$] and a main effect of group [$F(1,22) = 10.3, p = .004, \eta_p^2 = .319$], but no significant effect of modulation rate and no interactions (see Table 2A). Overall, amusics performed worse than controls at FM detection, and both groups performed better on the one-compared to the two-interval task. Similarly, the ANOVA for AM revealed a main effect of number of intervals [$F(1,22) = 48.5, p < .0001, \eta_p^2 = .688$] and a main effect

of group [$F(1,22) = 14.2, p = .001, \eta_p^2 = .393$]. There was also a main effect of modulation rate [$F(1,22) = 5.88, p = .024, \eta_p^2 = .211$], and none of the interactions were significant (see Table 2B). The main effect of modulation rate for AM is consistent with previous literature on sinusoidal AM detection with gated carriers (Moore and Sek, 1995; Sheft and Yost, 1990; Viemeister, 1979; Whiteford and Oxenham, 2015), and has been attributed to the increase in the number of cycles for faster rates. Typically, the opposite effect is observed with sinusoidal FM detection (Demany and Semal, 1989; Moore and Sek, 1996, 1994, 1992; Whiteford and Oxenham, 2015), with an increase in performance at slower rates relative to faster rates. Our findings of differential trends of modulation rate for AM (modest effect) versus FM (no effect) are consistent with previous studies.

Further analyses of response bias (c) (Fig. S2) and overall sensitivity between amusics and controls on one-interval FM and AM detection can be found in the supplementary materials (see S3.1. and S3.2.). Notably, the overall sensitivity to FM and AM was equal between amusics and controls when calculating one d' across all modulation indices (see Fig. S3), as intended by setting the differences in FM and AM depth individually.

3.2.2. Comparing one-interval slow FM detection with three-interval frequency discrimination—To further assess the potential effect of memory load on pitch discrimination, performance in the one-interval slow FM-detection task (low memory load) was compared to performance in the three-interval frequency-discrimination task from experiment 1 (high memory load). More specifically, the (log-transformed) peak-to-peak frequency difference at threshold in the one-interval FM task was compared to the (log-transformed) FDL in the discrimination task. Signal detection theory provides a way to compare performance based on the measure of sensitivity, d' , which should be independent of task. For the one-interval tasks, the modulation index that corresponded to a d' of 1.28 (the sensitivity equivalent to the 70.7% correct point for a two-interval AFC task) (Macmillan & Creelman, 2005, pg. 429) was estimated for each subject using linear regression. This was necessary to ensure that sensitivity for FM and sensitivity for AM corresponded to the same values of d' for both the one- and three-interval tasks. The slow rather than fast FM task was used because the detection of slow FM ($f_m < \sim 10$ Hz) is believed to rely on the same timing-based peripheral code as frequency discrimination of discrete tones (e.g., Sek and Moore, 1995). The task (FMDL vs. FDL) was the within-subjects factor, and group (amusic vs. control) was the between-subjects factor. All p values less than .05 were considered significant. Results indicated a main effect of task [$F(1,22) = 43.4, p < .0001, \eta_p^2 = .664$], a main effect of group [$F(1,22) = 12.1, p = .002, \eta_p^2 = .355$], and a significant interaction [$F(1,22) = 4.54, p = .045, \eta_p^2 = .171$]. Overall, slow FM thresholds were lower (better) than the FDLs; all subjects had better performance on the one-interval relative to the three-interval task (Fig. 4). Post-hoc tests using Bonferroni correction ($\alpha = .025$) indicated that amusics performed more poorly than controls on both measures [FDLs: $t(22) = -2.92, p = .004$, one-tailed; FMDLs: $t(22) = -3.06, p = .003$, one-tailed]. Therefore, the group by task interaction can be interpreted by viewing the differential trends in average performance for FDLs versus FMDLs in amusics and controls in Fig. 4. As is clear from Fig. 4, the degree of impairment was greater for the three-interval frequency

discrimination task compared to the one-interval FM-detection task in amusics relative to controls.

3.3. Discussion

The ability of amusics to detect FM in a one- or two-interval task was significantly poorer than that of controls at both slow ($f_m = 4$ Hz) and fast ($f_m = 20$ Hz) rates. This outcome supports the hypothesis that amusics have an underlying deficit in fine-grained frequency discrimination that extends beyond problems with memory or segmentation. However, the fact that the performance of amusics was equally degraded at FM rates of both 4 and 20 Hz is not, at face value, consistent with the suggestion that amusia results in particularly poor processing of rapid fine-grained pitch information (Albouy et al., 2016). If this were the case, one might expect an interaction between group and modulation rate, with poorer thresholds in amusics at faster modulation rates. It is possible, however, that the task design in the present study is too different from Albouy et al. to expect generalization, perhaps because the tones here are much longer and the changes over time are continuous, rather than discrete. Either poor rapid pitch encoding in amusics only applies to steady tones, and not FM, or the interaction Albouy et al. (2016) observed between group and tone duration reflected ceiling effects in the control group at the longer durations.

Surprisingly, AM detection thresholds were also worse in amusics relative to controls. Despite the fact that the task had low memory load requirements and did not involve any pitch-related changes, the amusics were still at a disadvantage. Tillmann et al. (2016) recently reported impaired reaction times for large intensity changes (20 dB) in amusics relative to controls, but these impairments were only present when gaps between tones were longer (1500 ms) as opposed to shorter (500 ms). These results suggest that the impaired reaction times found in Tillmann et al. could be a result of poorer intensity encoding, perhaps creating a slightly weaker memory trace even for large intensity changes. The present results of poor AM detection, however, was generally not expected, given that amusia is believed to be a pitch-specific deficit.

Trends for FM and AM detection were consistent in the two- and one-interval detection tasks. Overall, amusics performed more poorly at FM and AM detection tasks relative to controls, and their performance was not differentially affected by one- versus two-intervals. The lack of interaction between group and number of intervals for FM (or AM) detection was not unexpected, as the two-interval task could be completed using the same strategy as the one-interval task by, for instance, attending to just the first of the two intervals and determining whether or not it contained a modulated tone. Better performance across all subjects for the one-versus the two-interval tasks was modest for FM and moderate for AM, and could be related to training effects rather than memory, as the two-interval tasks always preceded the one-interval measures. Therefore, it is possible that learning occurred in both groups between the two modulation measures. For example, He et al. (2007) found the opposite effect for FMDLs, with better performance on three- vs. one-interval tasks, and they measured FMDLs in the opposite order as the present study (i.e., one-interval first). The general consensus is that amusia is a lifelong deficit, however, and performance on pitch related tasks cannot improve with training (Hyde and Peretz, 2004). This makes the ordering

effect hypothesis unlikely, as there was no interaction between group and number of intervals or task type and number of intervals. It is also possible that one-interval tasks are comparatively easier for our relatively untrained listeners because there are fewer possible strategies that the listener could adopt or switch between, or because of an overall lighter memory load. Further research is needed to clarify this potential discrepancy.

Comparisons of one-interval FMDLs with three-interval FDLs indicated that both amusics and controls were worse in the three-interval discrimination task than would be predicted by performance in the one-interval FM detection task. This difference may be due to the additional memory load required in the three-interval task. Most importantly, however, the detrimental effect of three intervals, relative to one, was markedly greater for the amusics than for the control participants: with the three-interval task, amusics' performance was worse by a factor of 8.26 (mean FDL = 7.71%; mean FMDL = 0.934%), whereas controls were worse by only a factor of 2.94 (mean FDL = 1.88%; mean FMDL = 0.638%). In addition, the effect size of group for the one- and two-interval FM tasks was substantially smaller than that observed for the three-interval task. Thus, it seems that amusia reflects poorer basic coding of frequency (and amplitude), as well as a poorer ability to compare frequencies across time. Experiment 3 explores the latter deficit in more detail.

4. Experiment 3: Comparing discrete and continuous changes in frequency

One potential explanation for why amusics seem more severely affected by having to compare frequencies across time is that their memory trace degrades more rapidly over time than it does for normal controls. Another potential explanation relates to the fact that the changes in pitch in the three-interval task were discrete, or segregated, rather than continuous. Indeed, Foxton et al. (2004) found that for both amusics and controls, the ability to detect a change in pitch was worse for segmented versus gliding tones; however, the segmented tones had 100 ms gaps of silence between them, potentially increasing memory load and/or decreasing pitch salience. Liu et al. (2015) found that Mandarin speaking amusics had deficits in segmented speech processing in speech with a flattened F0 (no pitch changes), suggesting that processing segmented sequences may be impaired in amusia, independent of pitch processing. It is therefore possible that segmentation, rather than memory decay *per se*, may be responsible for this deficit. In the general population, the introduction of segmentation tends to elevate (worsen) pitch-discrimination and pitch-change detection thresholds (Demany et al., 2009; Lyzenga et al., 2004; Sek and Moore, 1999), perhaps because comparing pitch changes between two auditory objects is more difficult than tracking changes within one (continuous) object (Demany et al., 2009).

We attempted to distinguish between explanations based on memory decay and segmentation by measuring performance in a two-interval task that involved detecting discrete frequency changes, but using the same time relationships as were used in the two-interval FM detection task of experiment 2 (see Fig. 5). If the amusic deficit is due at least in part to difficulties produced by segmentation, then performance in the task with discrete changes should be worse than in the FM-detection task with continuous changes. On the other hand, if the amusic deficit is only due to a more rapid decay in memory regardless of segmentation, then performance in the two tasks should be similar, as they share the same

overall duration. Note that a segmentation-related deficit in amusia would not rule out the memory decay hypothesis, as these two interpretations are not necessarily mutually exclusive. Segmentation could, theoretically, put a greater strain on memory resources (i.e., by creating the percept of multiple objects), even when the duration of the stimuli is held constant.

4.1. Methods

4.1.1. Participants—Twenty subjects (10 amusics) participated in experiment 3. One of the amusics had not previously participated in experiments 1 or 2, while all other subjects had completed the first two studies. Subjects were matched on the same demographic variables as in experiments 1 and 2 (see Table 3). All except one of the amusics performed below the average cutoff of the pitch-based subtasks on the MBEA. Participants provided written informed consent and were compensated with course credit or hourly payment for their time.

Audiometric thresholds were measured at octave frequencies from 250 to 8000 Hz. All but one participant had normal hearing up through 8 kHz, defined as audiometric thresholds no higher than 20 dB HL. One participant from the control group did not meet this criterion but had normal hearing up through 4 kHz and a small loss at 8 kHz (right ear: 25 dB HL; left ear: 35 dB HL). A t-test (two-tailed) revealed no difference in the 8-kHz audiometric thresholds (averaged across ears) between the groups ($t(22) = -.986, p = .337$). It should be noted that the controls had a slightly but significantly poorer four-tone pure-tone average (average of audiometric thresholds between 500 and 4000 Hz) (Table 3).

4.1.2. Stimuli and procedures—As in experiments 1 and 2, experiment 3 was administered in a sound-attenuating chamber via Sennheiser HD650 headphones at 60 dB SPL. All stimuli were generated in MATLAB using a 24-bit L22 soundcard (LynxStudio, Costa Mesa, CA) at a sampling rate of 48,000 Hz.

Detection of the segmented frequency changes was measured using a 2-interval, 2AFC adaptive procedure. The stimuli and procedures were designed to be analogous to the design from the two-interval, slow ($f_m = 4$ Hz) FM adaptive procedure in experiment 1. Thus, the target sequence contained frequency changes centered around 500 Hz, but the changes were discrete, rather than continuous. Both the reference and the target sequences were 1.5 s in duration and were comprised of 12 pure tones of 125 ms duration, alternating in frequency between $(500 + f)$ and $(500 - f)$ Hz. The reference sequence consisted of 12 sequential 500-Hz pure tones, also of 125 ms duration. Each individual tone within a sequence (target or reference) had 31.25-ms raised cosine onset and offset ramps, and there were no gaps in between tones within a sequence. The two sequences were separated from each other by a 500-ms gap. Subjects were instructed that they would hear two sequences of tones, one at a time, and their task was to select the sequence that sounded as if it were “changing.” Participants were reminded to attend to the feedback after each trial. All other aspects of the design and procedures were identical to the two-interval FM task from experiment 2. The new participant also completed the two-interval FM detection task for $f_m = 4$ Hz described in experiment 2.

4.2. Results

4.2.1. Discrete versus continuous FM detection—Results from the experiment are shown in Fig. 6, along with the replotted thresholds from the 2-interval 4-Hz FM detection task of experiment 2. A mixed-model ANOVA was performed on the log-transformed thresholds, with group (amusic or control) as an across-subjects factor and task type (segmented tones from this experiment or two-interval 4-Hz FM detection from experiment 2) as a within-subjects factor. For the purposes of this statistical analysis, the one new subject was treated as the direct replacement for the subject from experiment 1 who did not complete experiment 3. All p values less than .05 were considered significant. Consistent with experiment 2, there was a main effect of group [$F(1,18) = 8.41$, $p = .01$, $\eta_p^2 = .319$], with amusics performing significantly worse than controls. There was also a main effect of task type [$F(1,18) = 24.0$, $p = .0001$, $\eta_p^2 = .571$] and a significant interaction [$F(1,18) = 7.05$, $p = .016$, $\eta_p^2 = .281$]. All participants were worse at detecting frequency changes in segmented sequences compared to continuous changes in frequency (FM). Independent samples t tests indicated that amusics were significantly worse at continuous [$t(18) = -1.94$, $p = .034$, one-tailed] and segmented [$t(18) = -3.06$, $p = .004$, one-tailed] FM detection relative to controls, but the degree of impairment was greater in the segmented task (see Fig. 6). Because the segmented FM task was always measured after the continuous FM task, the current results may even underestimate the cost of segmentation for amusics if there was any learning between tasks.

4.3. Discussion

The results from experiment 3 indicate that segmentation led to poorer performance for both groups, but the effect was more detrimental to amusics' thresholds than to those of the controls. Importantly, experiment 3 controlled for the number of intervals (2AFC) and tone duration (1.5 s) by keeping these factors constant between the continuous and discrete FM tasks. The results indicate that segmentation contributes to poor pitch perception in amusia, and the effect of segmentation cannot be accounted for by tone duration.

5. General Discussion

The purpose of this study was to provide insights into the mechanisms underlying the deficits in pitch perception observed in people with congenital or developmental amusia using well-controlled psychophysical paradigms. A number of previous studies have assessed pitch perception in amusia — according to Vuvan et al. (2015), as of November 2013, 43 unique articles have addressed this topic. However, none have assessed the perception of high frequencies, where both frequency discrimination (Micheyl et al., 2012; Moore, 1973) and melodic perception (Attneave and Olson, 1971; Oxenham et al., 2011) is poor under normal conditions. Furthermore, no studies have assessed how different methodological paradigms using a low or high memory load (e.g., one-interval vs. three-interval AFC) differentially affect pitch perception thresholds in amusics versus controls.

The main findings of this study were: (1) Amusic deficits in pitch discrimination persist even at very high frequencies (8 kHz), despite the lack of musical pitch perception at these high frequencies, even in normal individuals (experiment 1); (2) amusics have deficits in FM

perception, suggesting that the deficit involves the representation and coding of frequency, rather than simply memory deficits (experiment 2); (3) the deficits in FM extend to AM as well, suggesting that amusia may involve deficits in fine-grained perception of dimensions other than frequency and pitch (experiment 2); and (4) the impaired ability to compare frequencies over time is not just a function of a more rapid memory decay over time, but appears to reflect a difficulty in comparing information across different auditory “objects” (experiment 3). Note that this effect of segmentation likely affects the nature of how the stimulus memory is stored, so memory decay and segmentation are unlikely to be mutually exclusive.

5.1. Poor high-frequency pitch perception

Results showed that amusics had poor pitch perception at low (500 and 2000 Hz) and high (8000 Hz) pure-tone frequencies. Poor perception at musically relevant frequencies coincides with classic findings suggesting that amusia is characterized by an underlying deficit in fine-grained pitch perception (e.g., Foxton et al., 2004; Peretz et al., 2002), with thresholds on the order of a semitone, depending on the particular paradigm used. Our findings suggest that the extent of the pitch perception deficit in part depends on the frequency where pitch perception is assessed, with thresholds significantly higher in both amusics and controls at 8000 Hz relative to the lower frequencies. Poor perception for high pure-tone frequencies in normal-hearing listeners is typically attributed to the use of a less-precise, rate-place code for frequency in the peripheral auditory system (Moore, 1973; Moore and Ernst, 2012). The main effect of frequency and the lack of group by frequency interaction we observed are consistent with this hypothesis. However, the main effect of group, with amusics nearly a factor of 2 worse at all frequencies relative to controls, suggests that central processing can also greatly limit pitch perception. The phenomenon that pitch perception is worse in amusics at 8 kHz, coupled together with recent findings of poor perception for resolved but not unresolved harmonics (Cousineau et al., 2015), supports the hypothesis that amusia is a deficit in central processing specific to spectro-temporal fine structure associated with coding pure tones and resolved harmonics.

The large inter-individual variance (see Fig. 1A) demonstrates that not all amusics have poor pitch perception, even when all but one of the amusics from experiment 1 were below cutoff on the pitch-based sub-tasks on the MBEA. This trend has been observed in numerous previous studies that have used adaptive paradigms (e.g., Foxton et al., 2004), reflecting the heterogeneity of the disorder (Vuvan et al., 2015). It is this heterogeneity, and the reliance of short-term or working memory in many pitch perception paradigms, that likely accounts for the possible effects of memory and/or segmentation observed in our sample. Indeed, the largest effect size of group was observed in the three-interval frequency discrimination task, where there were both memory and segmentation demands, while the smallest effect size of group was observed for one-interval continuous FM detection task, where the memory demands were minimized.

5.2. Deficits in one-interval, FM detection

One way to control for possible memory confounds in multiple-interval pitch perception paradigms, which require the comparison of pitch information across time, is to use one-

interval paradigms. This is because the use of detection paradigms requires simply noticing a change in frequency, rather than comparing information across time or comparing information to a built-in, memory representation (e.g., Jesteadt and Sims, 1975). Even when controlling for possible memory and segmentation confounds, amusics appeared to have impaired pitch perception (measured via slow-rate and fast-rate FM detection) relative to controls. This indicates that amusics have an underlying pitch and frequency coding deficit that cannot be explained entirely by memory load or segmentation, but this effect is much smaller than observed in more traditional frequency discrimination paradigms (e.g., experiment 1). Furthermore, our results could suggest amusia is not a problem with processing rapid, fine-grained pitch information, as proposed by Albouy et al. (2016). This is because amusics were equally impaired at processing slow-rate vs. fast-rate FM. Albouy et al. (2016) found that amusics were impaired at detecting pitch changes in short (100 ms) but not long (350 ms) duration tones relative to controls. While the present study only used long duration FM tones (1500 ms), potentially limiting the comparison to Albouy et al. (2016), a given cycle of FM is much longer for slow-rate FM (250 ms when $f_m = 4$ Hz) compared to fast-rate FM (50 ms when $f_m = 20$ Hz). Findings from Albouy et al. (2016) can possibly be attributed to using constant stimuli methods, where all participants were presented with the same f (1 or 2 semitones), leading to near ceiling pitch-change detection performance in the controls for the longer-duration tones. However, it is also possible that the mechanism for encoding pitch is different for fast-rate FM compared to slow-rate FM (e.g., Demany and Semal, 1989; Moore and Sek, 1996, 1995, 1994, 1992) and low frequency ($< \sim 4\text{--}5$ kHz) steady tones (Moore, 1973; Moore and Ernst, 2012), complicating the comparison between the present study and Albouy et al. (2016).

5.3. Impairment in one-interval AM detection

Interestingly, amusics also had poorer AM detection thresholds than controls, which has not been previously observed. Cousineau et al. (2012) found that amusics could not perceive a difference between consonant and dissonant intervals, and that this deficit was due to an impairment in the perception of harmonicity (i.e., the frequency spacing of harmonics at integer multiples of F_0 in a complex tone) but not amplitude modulated beats (i.e., a perception of “roughness” that occurs when two sinusoidal components fall within the same auditory filter). However, in Cousineau et al.’s experiment, the AM depth was always 100%, which was well above the threshold for amusics. Experiment 2 demonstrated poor AM detection when the changes in AM were more fine-grained, or near threshold with lower detectability — a considerably different process from AM at highly detectable depths, such as that used by Cousineau et al. (2012). A recent study examining slow ($f_m = 1$ Hz) and fast ($f_m = 20$ Hz) FM and AM detection at 500 Hz in a large cohort of young, normal-hearing listeners found that detection for near-threshold FM and AM was highly correlated, even across modulation type and modulation rate (Whiteford and Oxenham, 2015). One possible explanation for high multi-collinearity in modulation detection is that near-threshold FM and AM may use a shared, cortical rate-place population code (Micheyl et al., 2013). Micheyl et al. (2013) used mathematical modeling to combine human psychophysical data with animal physiological data and demonstrated that it was possible that the population of cortical neurons that code fine-grained changes in frequency may also code fine-grained changes in intensity. Evidence so far, however comes from frequency and intensity discrimination

thresholds, and has not yet been applied to modulation detection tasks. If poor FM and AM in amusia is due to a shared cortical code, then amusics should also exhibit poor intensity discrimination in a paradigm similar to that used in experiment 1.

5.4. Auditory Object Perception

Pitch discrimination was poorer in amusics than in controls, whether the task involved high or low memory load or segregated versus continuous changes in frequency. However, the magnitude of the deficit varied substantially, with amusics performing most poorly in tasks with the highest memory load and segregated tones and performing best in tasks with the lowest memory load and continuous changes in frequency. To illustrate this difference, the average thresholds for amusics were a factor of 8.26 better (lower) in the one-interval FM detection task than in the three-interval frequency-discrimination task. Control participants also had lower thresholds in the one-interval FM task than in the three-interval frequency-discrimination task, but the improvement was only by a factor of 2.94 on average.

Data from amusic participants also illustrated a multiple-interval effect for non-pitch tasks, as illustrated by the one- versus two-interval AM detection, with worse performance on the two-interval task. However, the decrease in performance for the two-interval task relative to the one-interval task was no greater than that observed in the control group. Analogous results were found for the one- versus two-interval FM task. Because the two-interval modulation-detection tasks could be performed using the same strategy as the one-interval task (i.e., by just attending to the second tone in a given trial), it is unclear exactly why performance was elevated across all subjects on the 2-interval task. It could be that participants adopted a poorer listening strategy for 2AFC, imparting memory and segmentation demands when they were not necessary. Alternatively, the relatively better performance of participants in both groups in the one-interval tasks may be result of practice, as the two-interval task was always completed first. However, earlier attempts to train amusics on pitch and melody-related tasks have been unsuccessful (Hyde and Peretz, 2004; Mignault Goulet et al., 2012), casting some doubt on the potential effects of practice, particularly for the amusic group.

Matched controls in the present study and non-amusic individuals in previous studies also exhibit poorer pitch perception in the context of segmented versus non-segmented tones (Demany et al., 2009; Lyzenga et al., 2004; Sek and Moore, 1999). It is not clear what mechanism may be responsible for poor perception of segmented tones, although it is likely separate from pitch processing. Previous studies have suggested the presence of a frequency-change detection mechanism, active only in the presence of continuous frequency changes, but not active during the detection of segmented frequencies (Lyzenga et al., 2004; Sek and Moore, 1999). However, Demany et al. (2009) found that participants were worse at detecting continuous frequency changes that were segmented via changes in amplitude (i.e., the envelope was segmented) compared to the same continuous changes without amplitude segmentation. Their results demonstrate that segmentation effects are not specific to frequency coding, and instead suggest that segmentation could create the percept of multiple auditory objects, which in turn may require greater perceptual or attentional load to process than one continuous, auditory object.

5.5. Implications on the nature of amusia

The present study suggests multiple auditory deficits in amusia that extend beyond poor fine-grained pitch perception. However, results from all three experiments show extensive overlap between amusics and controls (Figs. 1A, S1, 4A, and 6A), showing that not all amusics have the same kinds of underlying psychophysical deficits, adding further evidence to the heterogeneity of the disorder (e.g., Vuvan et al., 2015). All amusics do, however, share the commonality of poor melody perception, as assessed with the MBEA. The combined findings suggest underlying perceptual deficits in amusics are typically small and vary across individuals. These issues become further compounded with added demands of memory and segmentation, and hence become quite apparent in more complex tasks, such as musical listening.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Jiao Xu, Lindsey Ma, and Jeffrey Boucher for assistance with data collection. We also thank Anne Caclin for helpful comments and suggestions.

Funding: This work was supported by the National Institute of Health [grant R01 DC005216].

References

- Albouy P, Cousineau M, Caclin A, Tillmann B, Peretz I. Impaired encoding of rapid pitch information underlies perception and memory deficits in congenital amusia. *Sci Rep*. 2016; 6:18861.doi: 10.1038/srep18861 [PubMed: 26732511]
- Albouy P, Mattout J, Bouet R, Maby E, Sanchez G, Aguera PE, Daligault S, Delpuech C, Bertrand O, Caclin A, Tillmann B. Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain*. 2013a; 136:1639–1661. DOI: 10.1093/brain/awt082 [PubMed: 23616587]
- Albouy P, Mattout J, Sanchez G, Tillmann B, Caclin A. Altered retrieval of melodic information in congenital amusia: insights from dynamic causal modeling of MEG data. *Front Hum Neurosci*. 2015; 9:1–13. DOI: 10.3389/fnhum.2015.00020 [PubMed: 25653611]
- Albouy P, Schulze K, Caclin A, Tillmann B. Does tonality boost short-term memory in congenital amusia? *Brain Res*. 2013b; 1537:224–232. DOI: 10.1016/j.brainres.2013.09.003 [PubMed: 24041778]
- ANSI. American national standard acoustical terminology. New York: American National Standards Institute; 2013.
- Arehart KH. Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners. *J Acoust Soc Am*. 1994; 95:3574–3585. DOI: 10.1121/1.409975 [PubMed: 8046147]
- Attneave F, Olson RK. Pitch as a medium: A new approach to psychophysical scaling. *Am J Psychol*. 1971; 84:147–166. [PubMed: 5566581]
- Ayotte J, Peretz I, Hyde K. Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain*. 2002; 125:238–251. DOI: 10.1093/brain/awf028 [PubMed: 11844725]
- Bernstein JG, Oxenham AJ. Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number? *J Acoust Soc Am*. 2003; 113:3323–3334. DOI: 10.1121/1.1572146 [PubMed: 12822804]

- Carcagno S, Plack CJ. Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol*. 2011; 12:89–100. DOI: 10.1007/s10162-010-0236-1 [PubMed: 20878201]
- Carlyon RP, Shackleton TM. Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms? *J Acoust Soc Am*. 1994; 95:3541–3554. DOI: 10.1121/1.409971
- Chen JL, Kumar S, Wiliamson V, Scholz J, Griffiths T, Stewart L. Detection of the arcuate fasciculus in congenital amusia is dependent on tractography algorithm. *Proc 20th Annu Meet Organ Hum Brain Mapp*. 2015; 6:1–11. DOI: 10.3389/fpsyg.2015.00009
- Cousineau M, McDermott JH, Peretz I. The basis of musical consonance as revealed by congenital amusia. *Proc Natl Acad Sci U S A*. 2012; 109:19858–63. doi:10.1073/pnas.1207989109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1207989109. [PubMed: 23150582]
- Cousineau M, Oxenham AJ, Peretz I. Congenital amusia: A cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia*. 2015; 66:293–301. DOI: 10.1016/j.neuropsychologia.2014.11.031 [PubMed: 25433224]
- Demany L, Carlyon RP, Semal C. Continuous versus discrete frequency changes: different detection mechanisms? *J Acoust Soc Am*. 2009; 125:1082–1090. DOI: 10.1121/1.3050271 [PubMed: 19206883]
- Demany L, Semal C. Detection thresholds for sinusoidal frequency modulation. *J Acoust Soc Am*. 1989; 85:1295–1301. DOI: 10.1121/1.397460 [PubMed: 2708671]
- Durlach NI, Braida LD. Intensity perception. I. Preliminary theory of intensity resolution. *J Acoust Soc Am*. 1969; 46:372–383. DOI: 10.1121/1.1911699 [PubMed: 5804107]
- Foxton JM, Dean JL, Gee R, Peretz I, Griffiths TD. Characterization of deficits in pitch perception underlying “tone deafness. *Brain*. 2004; 127:801–810. [PubMed: 14985262]
- Gaser C, Schlaug G. Brain structures differ between musicians and non-musicians. *J Neurosci*. 2003; 23:9240–9245. doi:23/27/9240[pil]. [PubMed: 14534258]
- Gosselin N, Jolicœur P, Peretz I. Impaired memory for pitch in congenital amusia. *Ann N Y Acad Sci*. 2009; 1169:270–272. DOI: 10.1111/j.1749-6632.2009.04762.x [PubMed: 19673791]
- Houtsma AJM, Smurzynski J. Pitch identification and discrimination for complex tones with many harmonics. *J Acoust Soc Am*. 1990; 87:304.doi: 10.1121/1.399297
- Hyde KL, Lerch JP, Zatorre RJ, Griffiths TD, Evans AC, Peretz I. Cortical thickness in congenital amusia: when less is better than more. *J Neurosci*. 2007; 27:13028–13032. DOI: 10.1523/JNEUROSCI.3039-07.2007 [PubMed: 18032676]
- Hyde KL, Peretz I. Brains that are out of tune but in time. *Psychol Sci*. 2004; 15:356–360. DOI: 10.1111/j.0956-7976.2004.00683.x [PubMed: 15102148]
- Hyde KL, Zatorre RJ, Griffiths TD, Lerch JP, Peretz I. Morphometry of the amusic brain: A two-site study. *Brain*. 2006; 129:2562–2570. DOI: 10.1093/brain/awl204 [PubMed: 16931534]
- Hyde KL, Zatorre RJ, Peretz I. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cereb Cortex*. 2011; 21:292–299. DOI: 10.1093/cercor/bhq094 [PubMed: 20494966]
- Jesteadt W, Sims SLS. Decision processes in frequency discrimination. *J Acoust Soc Am*. 1975; 57:1161–1168. [PubMed: 1127170]
- Jiang C, Lim VK, Wang H, Hamm JP. Difficulties with Pitch Discrimination Influences Pitch Memory Performance: Evidence from Congenital Amusia. *PLoS One*. 2013; 8:1–14. DOI: 10.1371/journal.pone.0079216
- Joris, PX., Verschooten, E. Basic Aspects of Hearing: Physiology and Perception. Springer; New York: 2013. On the limit of neural phase locking to fine structure in humans; p. 101-108.
- Lévesque Y, Fauvel B, Groussard M, Caclin A, Albouy P, Platel H, Tillmann B. Altered intrinsic connectivity of the auditory cortex in congenital amusia. *J Neurophysiol*. 2016; 116:88–97. DOI: 10.1152/jn.00663.2015 [PubMed: 27009161]
- Levitt H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am*. 1971; 49:467–477. DOI: 10.1121/1.1912375
- Loui P, Alsop D, Schlaug G. Tone deafness: a new disconnection syndrome? *J Neurosci*. 2009; 29:10215–10220. DOI: 10.1523/JNEUROSCI.1701-09.2009 [PubMed: 19692596]

- Lyzenga J, Carlyon RP, Moore BCJ. The effects of real and illusory glides on pure-tone frequency discrimination. *J Acoust Soc Am*. 2004; 116:491–501. DOI: 10.1121/1.1756616 [PubMed: 15296008]
- Macmillan, NA., Creelman, CD. *Detection Theory: A User's Guide*. 2nd. Lawrence Erlbaum Associates, Inc; 2005.
- Micheyl C, Delhommeau K, Perrot X, Oxenham AJ. Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res*. 2006; 219:36–47. DOI: 10.1016/j.heares.2006.05.004 [PubMed: 16839723]
- Micheyl C, Oxenham AJ. Sequential F0 comparisons between resolved and unresolved harmonics: no evidence for translation noise between two pitch mechanisms. *J Acoust Soc Am*. 2004; 116:3038–3050. DOI: 10.1121/1.1806825 [PubMed: 15603149]
- Micheyl C, Schrater PR, Oxenham AJ. Auditory frequency and intensity discrimination explained using a cortical population rate code. *PLoS Comput Biol*. 2013; 9:e1003336.doi: 10.1371/journal.pcbi.1003336 [PubMed: 24244142]
- Micheyl C, Xiao L, Oxenham AJ. Characterizing the dependence of pure-tone frequency difference limens on frequency, duration, and level. *Hear Res*. 2012; 292:1–13. DOI: 10.1016/j.heares.2012.07.004 [PubMed: 22841571]
- Mignault Goulet G, Moreau P, Robitaille N, Peretz I. Congenital amusia persists in the developing brain after daily music listening. *PLoS One*. 2012; 7doi: 10.1371/journal.pone.0036860
- Moore BCJ. Frequency difference limens for short-duration tones. *J Acoust Soc Am*. 1973; 54:610–619. [PubMed: 4754385]
- Moore BCJ, Ernst SMA. Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code. *J Acoust Soc Am*. 2012; 132:1542–1547. DOI: 10.1121/1.4739444 [PubMed: 22978883]
- Moore BCJ, Moore Ga. Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects. *Hear Res*. 2003; 182:153–163. DOI: 10.1016/S0378-5955(03)00191-6 [PubMed: 12948610]
- Moore BCJ, Peters RW. Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity. *J Acoust Soc Am*. 1992; 91:2881–2893. DOI: 10.1121/1.402925 [PubMed: 1629481]
- Moore BCJ, Sek A. Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking. *J Acoust Soc Am*. 1996; 100:2320–2331. DOI: 10.1121/1.417941 [PubMed: 8865639]
- Moore BCJ, Sek A. Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation). *J Acoust Soc Am*. 1995; 97:2468–2478. DOI: 10.1121/1.411967 [PubMed: 7714263]
- Moore BCJ, Sek A. Effects of carrier frequency and background noise on the detection of mixed modulation. *J Acoust Soc Am*. 1994; 96:741–751. [PubMed: 7930075]
- Moore BCJ, Sek A. Detection of combined frequency and amplitude modulation. *J Acoust Soc Am*. 1992; 92:3119–3131. DOI: 10.1121/1.404208 [PubMed: 1474227]
- Norman-Haignere SV, Albouy P, Caclin A, McDermott JH, Kanwisher NG, Tillmann B. Pitch-Responsive Cortical Regions in Congenital Amusia. *J Neurosci*. 2016; 36:2986–2994. DOI: 10.1523/JNEUROSCI.2705-15.2016 [PubMed: 26961952]
- Oxenham AJ, Micheyl C, Keebler MV, Loper A, Santurette S. Pitch perception beyond the traditional existence region of pitch. *Proc Natl Acad Sci U S A*. 2011; 108:7629–7634. DOI: 10.1073/pnas.1015291108 [PubMed: 21502495]
- Peretz I. Brain specialization for music: New evidence from congenital amusia. *Neuroscientist*. 2001; 930:153–165. DOI: 10.1177/107385840200800412
- Peretz I, Ayotte J, Zatorre RJ, Mehler J, Ahad P, Penhune VB, Jutras B. Congenital Amusia: A disorder of fine-grained pitch discrimination. *Neuron*. 2002; 33:185–191. DOI: 10.1016/S0896-6273(01)00580-3 [PubMed: 11804567]
- Peretz I, Champod AS, Hyde K. Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Ann N Y Acad Sci*. 2003; 999:58–75. [PubMed: 14681118]

- Peretz I, Cummings S, Dubé MP. The genetics of congenital amusia (tone deafness): a family-aggregation study. *Am J Hum Genet.* 2007; 81:582–588. DOI: 10.1086/521337 [PubMed: 17701903]
- Peretz I, Gosselin N, Tillmann B, Cuddy LL, Gagnon B, Trimmer CG, Paquette S, Bouchard B. On-line identification of congenital amusia. *Music Percept.* 2008; 25:331–344.
- Peretz I, Kolinsky R, Tramo M, Labrecque R, Hublet C, Demeurisse G, Belleville S. Functional Dissociations following Bilateral Lesions of Auditory Cortex. *Brain.* 1994; 117:1283–1301. <http://dx.doi.org/10.1093/brain/117.6.1283>. [PubMed: 7820566]
- Plack, CJ., Oxenham, AJ. The psychophysics of pitch. In: Plack, CJ, Oxenham, AJ, Fay, RR., Popper, AN., editors. *Pitch: Neural Coding and Perception*. Springer; New York: 2005. p. 7-55.
- Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci.* 2002; 5:688–694. DOI: 10.1038/nn871 [PubMed: 12068300]
- Sek A, Moore BCJ. Frequency discrimination as a function of frequency, measured in several ways. *J Acoust Soc Am.* 1995; 97:2479–2486. DOI: 10.1121/1.411968 [PubMed: 7714264]
- Sek A, Moore BC. Discrimination of frequency steps linked by glides of various durations. *J Acoust Soc Am.* 1999; 106:351–359. DOI: 10.1121/1.427061 [PubMed: 10420627]
- Sheft S, Yost WA. Temporal integration in amplitude modulation detection. *J Acoust Soc Am.* 1990; 88:796–805. DOI: 10.1121/1.399729 [PubMed: 2212305]
- Studebaker GA. A “rationalized” arcsine transform. *J Speech Hear Res.* 1985; 28:455–462. [PubMed: 4046587]
- Tillmann B, Lévêque Y, Fornoni L, Albouy P, Caclin A. Impaired short-term memory for pitch in congenital amusia. *Brain Res.* 2016; 1640:251–263. DOI: 10.1016/j.brainres.2015.10.035 [PubMed: 26505915]
- Tillmann B, Schulze K, Foxton JM. Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain Cogn.* 2009; 71:259–264. DOI: 10.1016/j.bandc.2009.08.003 [PubMed: 19762140]
- Viemeister NF. Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am.* 1979; 66:1364–1380. DOI: 10.1121/1.383531 [PubMed: 500975]
- Vuvan DT, Nunes-Silva M, Peretz I. Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex.* 2015; 69:186–200. DOI: 10.1016/j.cortex.2015.05.002 [PubMed: 26079675]
- Wechsler, D. Wechsler Abbreviated Scale of Intelligence. Second. Pearson; Bloomington, MN: 2011.
- Whiteford KL, Oxenham AJ. Using individual differences to test the role of temporal and place cues in coding frequency modulation. *J Acoust Soc Am.* 2015; 138:3093–3104. DOI: 10.1121/1.4935018 [PubMed: 26627783]
- Williamson VJ, McDonald C, Deutsch D, Griffiths TD, Stewart L. Faster decline of pitch memory over time in congenital amusia. *Adv Cogn Psychol.* 2010; 6:15–22. DOI: 10.2478/v10053-008-0073-5 [PubMed: 20689638]
- Williamson VJ, Stewart L. Memory for pitch in congenital amusia: beyond a finegrained pitch discrimination problem. *Memory.* 2010; 18:657–669. DOI: 10.1080/09658211.2010.501339 [PubMed: 20706954]
- Zhang Y, Moore DR, Guiraud J, Molloy K, Yan TT, Amitay S. Auditory Discrimination Learning: Role of Working Memory. *PLoS One.* 2016; 11:e0147320. doi: 10.1371/journal.pone.0147320 [PubMed: 26799068]

Highlights

- Poor pitch discrimination in amusia extends to high frequencies.
- Amusics have impaired frequency- and amplitude-modulation detection.
- Memory and segmentation demands exacerbate underlying auditory deficits in amusia.

Fig. 1.

Frequency discrimination at low and high frequencies in amusics and controls. Filled and open circles represent data from amusic and control participants, respectively. A. Individual FDLs as a function of frequency. B. Group mean FDLs as a function of frequency. Error bars represent ± 1 standard error of the mean. Amusic data points in (A) are offset horizontally for data visualization purposes.

Fig. 2.

Example psychometric function for one-interval FM detection for one subject. The x-axis corresponds to the peak-to-peak frequency change, expressed as a percentage of the carrier frequency. The y-axis corresponds to d' , calculated based on the hit and false-alarm rates for each of the modulation indices. Solid lines correspond to the regression line, fitted to the four data points for each modulation rate (blue: $f_m = 4$ Hz; black: $f_m = 20$ Hz). Threshold was defined as the peak-to-peak frequency change corresponding to a d' of 1.14 (dashed line).

Fig. 3. Average FM and AM thresholds for the two- versus one-interval tasks at slow ($f_m = 4$ Hz) and fast ($f_m = 20$ Hz) modulation rates. Closed circles represent amusic and open circles represent controls. The x-axis corresponds to modulation rate. The y-axis corresponds to the average (A–B) FM threshold, plotted as peak-to-peak frequency change as a percentage of the carrier frequency, and (C–D) AM threshold, plotted as modulation depth in logarithmic units. The left column plots thresholds for the two-interval tasks, while the right column plots the comparable thresholds for the one-interval tasks. Error bars represent ± 1 standard error of the geometric mean.

Fig. 4. Individual (A) and average (B) performance for three-interval FDLs vs. one-interval FMDLs. The x-axis represents task type, with three-interval frequency discrimination ($f = 500$ Hz; experiment 1) closest to the y-axis and one-interval FM detection ($f_c = 500$ Hz and $f_m = 4$ Hz; experiment 2) furthest from the y-axis. The y-axis represents percent frequency change at threshold, where the FMDLs were transformed to peak-to-peak frequency change to be comparable to the FDLs. Error bars represent ± 1 standard error of the geometric mean.

Fig. 5. Schematic diagram of target stimulus from experiments 2 and 3: Continuous (A) vs. discrete (B) FM. The x-axis corresponds to time, while the y-axis corresponds to frequency. The horizontal dashed line represents the carrier frequency, 500 Hz. The vertical dashed line represents the frequency excursion from the carrier (Δf), varied adaptively in the same manner for experiments 2 and 3. Blue lines correspond to the target stimulus. The frequency change contours of the target and reference stimuli were the only differences between experiments 2 and 3.

Fig. 6. Continuous vs. discrete slow ($f_m = 4$ Hz) FMDLs. Black circles represent amusics and white circles represent controls. The x-axis corresponds to type of modulation (continuous vs. discrete). The y-axis corresponds to the average FM threshold, plotted in peak-to-peak frequency change. Error bars represent ± 1 standard error of the geometric mean.

Table 1

Demographic averages for 12 amusics and 12 controls. Standard deviation is in parentheses. Pitch MBEA represents the average percent correct on the pitch-based subtasks of the MBEA. Global MBEA represents the overall average percent correct across all tasks on the MBEA. All MBEA percentages were transformed to rationalized arcsine units for statistical analyses (Studebaker, 1985). The 4 pure-tone average is the audiometric thresholds between 500–4000 Hz, averaged across frequencies and ears.

	Group		<i>p</i> -value
	Amusics (n = 12)	Controls (n = 12)	
Age (years)	23.42 (9.83)	23.33 (8.84)	.983
Musical Experience (years)	1.46 (1.53)	1.17 (2.08)	.7
Education (years)	15 (1.26)	15.13 (1.72)	.841
IQ	106 (9.94)	112.17 (10.03)	.144
Gender	9 females, 3 males	9 females, 3 males	–
Pitch MBEA (%)	65.83 (4.65)	85.46 (4.46)	< .0001*
Global MBEA (%)	67.5 (4.49)	86.3 (3.73)	< .0001*
Audiogram- 4 Pure-Tone Average (dB HL)	3.65 (3.07)	6.09 (3.72)	.093

All *p* values correspond to independent-samples *t* tests. The *p* values less than 0.05 are marked with an asterisk and shown in bold.

ANOVA results for comparing one- versus two-interval FM (A) and AM (B) detection thresholds. Group refers to amusics versus controls, Rate refers to slow ($f_m = 4$ Hz) versus fast ($f_m = 20$ Hz) modulation rates, and Int. to the number of intervals in a given task (one versus two).

Table 2

2A. FM Detection					
Effect	DFn	DFD	F	p	η_p^2
Group	1	22	10.3	.004*	.319
Rate	1	22	.013	.912	.001
Int.	1	22	9.34	.006*	.298
Group $\times$ Rate	1	22	1.26	.273	.054
Group $\times$ Int.	1	22	.317	.579	.014
Rate $\times$ Int.	1	22	.19	.667	.009
Group $\times$ Rate $\times$ Int.	1	22	.984	.332	.043
2B. AM Detection					
Effect	DFn	DFD	F	p	η_p^2
Group	1	22	14.3	.001*	.393
Rate	1	22	5.89	.024*	.211
Int.	1	22	48.5	< .0001*	.688
Group $\times$ Rate	1	22	.173	.681	.008
Group $\times$ Int.	1	22	.773	.389	.034
Rate $\times$ Int.	1	22	.047	.831	.002
Group $\times$ Rate $\times$ Int.	1	22	.825	.374	.036

Significant effects are bolded and marked with *.

Table 3

Demographic average data for the 10 amusic and 10 control participants from experiment 3. Standard deviations are shown in parentheses. Pitch MBEA scores represent the average percent correct on the pitch based subtasks of the MBEA. Global MBEA scores represent the overall average percent correct across all tasks on the MBEA. All MBEA percentages were transformed to rationalized arcsine units for statistical analyses (Studebaker, 1985). The 4 pure-tone average is the audiometric thresholds between 500–4000 Hz, averaged across frequencies and ears.

	Group		<i>p</i> -value
	Amusics (n = 10)	Controls (n = 10)	
Age (years)	23.6 (10.9)	23.3 (9.6)	.946
Musical Experience (years)	1.55 (1.61)	1.3 (2.26)	.779
Education (years)	14.9 (1.35)	14.9 (1.61)	> .999
IQ	106.7 (10.7)	113.3 (10.5)	.181
Gender	7 females, 3 males	7 females, 3 males	–
Pitch MBEA (%)	65.78 (4.25)	84.67 (4.41)	< .0001*
Global MBEA (%)	68.06 (4.63)	85.72 (3.66)	< .0001*
Audiogram- 4 pure-tone average (dB HL)	3.06 (2.29)	6.5 (3.75)	.024*

All statistical tests involved independent-samples *t* tests. Significant effects ($p < 0.05$) are bolded and marked with an asterisk.