SPEECH INTONATION PERCEPTION DEFICITS IN MUSICAL TONE DEAFNESS (CONGENITAL AMUSIA)

ANIRUDDH D. PATEL
The Neurosciences Institute, San Diego, California

MEREDITH WONG
The Neurosciences Institute and University of California, San Diego

JESSICA FOXTON
CerCo, UMR CNRS, Toulouse, France

ALIETTE LOCHY & ISABELLE PERETZ
Université de Montréal, Montréal, Canada

TO WHAT EXTENT DO MUSIC and language share neural mechanisms for processing pitch patterns? Musical tone-deafness (amusia) provides important evidence on this question. Amusics have problems with musical melody perception, yet early work suggested that they had no problems with the perception of speech intonation (Ayotte, Peretz, & Hyde, 2002). However, here we show that about 30% of amusics from independent studies (British and French-Canadian) have difficulty discriminating a statement from a question on the basis of a final pitch fall or rise. This suggests that pitch direction perception deficits in amusia (known from previous psychophysical work) can extend to speech. For British amusics, the direction deficit is related to the rate of change of the final pitch glide in statements/ questions, with increased discrimination difficulty when rates are relatively slow. These findings suggest that amusia provides a useful window on the neural relations between melodic processing in language and music.

Received October 2, 2007, accepted December 5, 2007.

Key words: speech intonation, melody, prosody, tone deafness, congenital amusia

MUSICAL TONE DEAFNESS OR 'CONGENITAL AMUSIA' (henceforth, amusia) has been noted for over a century (Allen, 1878). Recent years have seen a surge of research on this disorder aimed at understanding its nature and underlying causes. While many people in Western culture self-label as 'tone deaf,' most such individuals simply mean that they do not perceive themselves as good singers, rather than meaning that they have impaired perception of music (Cuddy, Balkwill, Peretz, & Holden, 2005). True amusia, in contrast, is distinguished by severe problems with music perception. For example, amusics have difficulty judging if two melodies are the same or different, in detecting when music is out of key (including their own singing), and in recognizing what should be familiar tunes from their culture (Ayotte et al., 2002). These problems cannot be attributed to hearing loss, lack of exposure to music, or to any obvious nonmusical social or cognitive impairments. The core deficit in this disorder concerns pitch processing (Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Hyde & Peretz, 2004). It appears that there is a genetic basis for this disorder (Drayna, Manichaikul, de Lange, Snieder, & Spector, 2001; Peretz, Cummings, & Dubé, 2007), and evidence from neuroimaging has revealed specific structural and functional differences between normal and amusic brains (Hyde, Lerch, Zatorre, Griffiths, Evans, & Peretz, 2007; Hyde, Zatorre, Griffiths, Lerch, & Peretz, 2006; Mandell, Schulze, & Schlaug, 2007; Peretz, Brattico, & Tervaniemi, 2005).[1]

From the standpoint of cognitive neuroscience, amusia is of interest because of its apparent specificity. Amusics often seem normal in every other way and can excel in other domains (for example, the Nobel laureate Milton Friedman was amusic). Hence amusia offers the opportunity to study how a selective cognitive deficit emerges, tracing the causal links between genes, physiology, and cognition. From a practical perspective, research on amusia is also attractive because affected individuals (who are estimated to comprise about 4%

---

[1] Amusia as discussed in the present article ('tone deafness') should be conceptually distinguished from an even rarer perceptual disorder in which musical sounds are perceived as severely distorted in sound quality. Such 'dystimbric' amusics perceive music as noise (for example, a piano may sound like the banging of pots and pans) and often find music aversive, in contrast to tone-deaf amusics of the type discussed here (cf. Sacks, 2007).

Music Perception VOLUME 25, ISSUE 4, PP. 357–368, ISSN 0730-7829, ELECTRONIC ISSN 1533-8312 DOI:10.1525/MP.2008.25.4.357

of the population) can readily be recruited via a process of advertisement and careful screening using standardized musical tasks (Peretz, Champod, & Hyde, 2003).

In pursuing links between genes and cognition in this disorder, the true specificity of the deficit to music is of substantial theoretical importance. That is, while amusics have no other obvious disorders in everyday life, it is important to know if they have subtle nonmusical deficits that reveal themselves in controlled laboratory investigations (cf. Douglas & Bilkey, 2007). If so, the nature of these deficits will be important in constraining hypotheses about the physiological effects of the genes in question.

One obvious domain for comparative research is speech, and in particular, the perception of speech intonation. Speech intonation (or "speech melody" as it is often called by linguists) involves the structured use of pitch patterns to convey a variety of structural, affective, pragmatic, and attitudinal information (Bolinger, 1985; Ladd, 1996; Patel, 2008). Since amusics have severe problems with musical melody perception, it is natural to ask if their perception of speech intonation is normal or impaired. This question was first addressed by Ayotte et al. (2002) and Peretz et al. (2002), who examined the ability of amusics to discriminate between sentences on the basis of intonation contour. Two types of sentence pairs were tested. In "statement-question" pairs, the intonation contours differed at the end of the sentence (e.g., "He likes to drive fast cars" spoken with a pitch fall or rise on "cars"). In "focus-shift" pairs, the intonation contours differed within the sentence (e.g., "Go in front of the bank, I said" spoken with a salient pitch accent on "front" or "bank"). (NB: The sentences were in French, since the study was conducted with French speaking amusics in Montreal). Amusics had no problem discriminating the sentences. In contrast, they had difficulty discriminating tone sequence analogs of intonation patterns created from the sentences (cf. Patel, Peretz, Tramo, & Labrecque, 1998 for details on how these analogs were constructed). These results suggested that amusics were unimpaired in speech intonation perception and that their problems were limited to music-like materials.

Subsequent work has nuanced this picture. A study by Lochy, Hyde, Parisel, Van Hyfte, and Peretz (2004) was the first to suggest that some amusics may in fact have a problem with the perception of speech intonation. These researchers tested 11 French-Canadian amusics with stimuli and tasks very similar to those used by Ayotte et al. (2002). This time, however, amusics had difficulty discriminating statements from questions, a task that control participants found easy. In fact, amusics performed *worse* on linguistic statement-question discrimination than on tone sequence analogs of these sentences. When the group data were analyzed more closely, this result was found to be due to a subset of four amusics.

Lochy et al.'s (2004) findings differ from those of Ayotte et al. (2002) and are important for issues of domain specificity. Hence they naturally call for replication and explanation. The present study attempts to replicate and understand these findings by examining statement-question discrimination in a group of British amusics using stimuli and methods similar to those of Lochy et al. (except with English rather than French sentences). As shown below, Lochy et al.'s findings do indeed replicate. To better understand these findings, we exploit the fact that amusics with statement-question (henceforth, SQ) discrimination problems in language do not commit errors on every SQ item, but fail on certain items and succeed on others. Thus we compare acoustic features of the correctly vs. incorrectly discriminated items in order to derive some clues as to what might be behind the problem in linguistic SQ discrimination.

In particular, we focus on aspects of the critical pitch movement in linguistic SQ pairs. In the stimuli used in the current study (and in the study of Lochy et al., 2004), members of each linguistic SQ pair are acoustically identical until the final region of the sentence, where pitch either glides up or down. Amusics are known to have problems in discriminating the direction of pitch movements in nonlinguistic tone sequences, even when these movements are above their thresholds for pitch change detection. For example, using pure tone glides Foxton et al. (2004) found that amusics needed a pitch movement over 20 times the size in semitones (on average) as that needed by controls in order to judge direction correctly.[2] Hence it seems plausible that problems in linguistic SQ discrimination may reflect problems in perceiving the direction of pitch glides in speech, if such glides are relatively small in size. To test this hypothesis, we examined the sizes (in semitones, henceforth st) of final pitch falls and rises in SQ pairs that were discriminated correctly vs. incorrectly.

---

[2] Foxton et al. (2004) used a pitch direction task which required explicit labeling of a pitch movement as 'up' or 'down.' More recent psychophysical work using a different task that avoids explicit labels has confirmed a pitch direction deficit in amusia, but suggests that thresholds are about 5 times higher than those of controls, in semitones (Griffiths et al., 2007).

In addition to glide size, we also examined glide rate, which measures the 'velocity' (in semitones/s) of a pitch rise or fall. Glide rate is distinct from glide size because glides of the same size can differ in rate depending on the duration of the pitch excursion. The motivation for examining glide rate comes from research on intonation perception in normal individuals. Such research suggests that glide rate is relevant to listeners' judgments about intonational contrasts, including contrasts between statements and questions (Gósy & Terken, 1994; Niebuhr, 2003). Hence we examine glide rate in linguistic SQ items that were successfully vs. unsuccessfully discriminated by our amusic participants.

In addition to studying British amusics, the current study also reexamines the original data of Lochy et al. (2004) in order to analyze glide size and glide rate in correctly vs. incorrectly discriminated SQ pairs. Hence this paper is divided into two studies: study 1 examines data from British amusics, while study 2 examines data from French-Canadian amusics.

Study 1: British Amusics

Method

The data for the British amusics were collected as part of an earlier study of intonation perception in amusia (Patel, Foxton, & Griffiths, 2005). That study reported results for focus-shift (FS) sentences, whereas this study reports the results of SQ discrimination.

PARTICIPANTS

10 British amusics (8 women, mean age 58.3 years, *SD* = 12.0) free of neurological or psychiatric disorders participated in the current study. All but one were right-handed. Their musical deficit was confirmed using the Montreal Battery of Evaluation of Amusia (MBEA, Peretz et al., 2003). All had normal hearing in at least one ear, defined as a mean hearing level of 20 dB HL or less, measured by pure tone audiometry at 250, 500, and 1000 Hz (see Foxton et al., 2004 for details).

STIMULI

The materials for SQ discrimination consisted of 12 pairs of sentences uttered by a female native speaker of American English (average F0 across all sentences = 189.2 Hz). Sentences ranged between 3 and 12 syllables (average duration = 1.6 s). Members of each pair were lexically identical but differed in intonation contour. Specifically, the final region of the intonation contour—usually the last syllable or word—differed in whether the speaker produced an upward or downward pitch glide (for a question or statement, respectively). Sentences were constructed using a cross-splicing technique so that members of a pair were acoustically identical until the final region. The waveform of the final region was edited so that across members of a pair the timing of syllables was identical and the acoustic waveform amplitude and perceived intensity were roughly equal, leaving pitch as the only salient cue for discrimination. Table 1 lists the sentences and gives some of their acoustic characteristics.

TABLE 1. Sentences Used in SQ Discrimination with British Amusics.

		Size of Final Pitch Glide (st)		Rate of Final Pitch Glide (st/s)	
Sentence	Rate (syl/s)	S	Q	S	Q
He speaks French./?	2.7	−7.9	12.6	−54.5	81.2
She plays the flute./?	3.5	−8.9	12.7	−52.5	79.8
She forgot her book./?	4.1	−7.0	12.5	−48.1	129.7
He wants to leave now./?	3.9	−4.7	12.3	−23.4	70.7
He likes to drive fast cars./?	3.4	−8.6	11.7	−33.9	54.4
He works ten hours a day./?	4.1	−5.2	12.8	−24.3	64.3
Francis is at the restaurant./?	4.5	−5.1	7.2	−45.3	59.5
The telephone doesn't work./?	5.7	−5.6	12.9	−45.5	74.8
He has been in Paris for three months./?	4.4	−10.0	13.4	−32.4	80.2
The supermarket is closed on Sunday./?	5.0	−9.8	11.7	−27.5	37.4
He wants to buy a house next to the beach./?	4.9	−6.1	11.7	−44.1	97.2
She drinks three large cups of coffee every morning./?	4.8	−10.1	15.5	−22.0	52.3
Mean	**4.2**	**−7.4**	**12.3**	**−37.8**	**73.5**
SD	0.8	2.0	1.9	11.9	23.9

Note: The syllable(s) on which pitch glided up (for questions) or down (for statements) is underlined in column 1.

For each sentence, a nonlinguistic tone sequence analog was created by replacing each syllable with a tone whose pitch was fixed at the frequency midway between the highest and lowest F0 of the syllable (see Patel et al., 1998 for details). Discrete tones were used rather than gliding tones (which would mimic intonation contours more precisely) in order to make the stimuli more music-like. All tones had a complex frequency structure consisting of a fundamental and seven odd harmonics of decreasing amplitude, giving the analogs a clarinet-like quality. Because of the way the sentence pairs were constructed, members of each tone sequence pair were identical in terms of the duration and rhythm of tones, and differed in pitch only on the final tones of the sequence. (See Figure 3 of Patel et al., 1998 for an example pitch contour of a sentence and its corresponding nonlinguistic tone sequence).

PROCEDURE AND SCORING

From the original 12 pairs of sentences, a list of 32 pairs of test stimuli were created as follows: Each pair was presented in 'same' and 'different' configuration (yielding 24 pairs), and 8 of these pairs were pseudorandomly selected for repetition (yielding 32 pairs, 16 each in 'same' and 'different' configuration). The 32 linguistic SQ stimuli were combined with 32 linguistic focus-shift stimuli in a random order and divided into two blocks of 32 stimuli each. These blocks were presented for same-different discrimination (the current study concentrates on data from the SQ items).

A procedure identical to that described above was used for the tone sequence stimuli, which were presented in different blocks than the linguistic stimuli. All participants completed the blocks in a fixed order (speech first, then tones), but the blocks were not always run in the same session, as the tests were part of a larger study requiring multiple visits (Foxton et al., 2004). Within each block members of a pair were separated by 2 s and pairs were separated by 5 s. Participants were instructed to listen to each pair and indicate with a button press if the members of the pair sounded identical or if they sounded different in any way. Practice items were given before each block. Participants were tested in a quiet room in the School of Neurology at the University of Newcastle upon Tyne.

Performance was analyzed in terms of hits and false alarms, in accordance with the methods of Ayotte et al. (2002). A hit was defined as a 'different' pair classified as different, and a false alarm was defined as a 'same' pair classified as different. The percentage of correct responses was also analyzed for comparison with control data reported in another study that used exactly the same stimuli as the current study (Nicholson, Baum, Kilgour, Koh, Munhall, & Cuddy, 2003).

IDENTIFYING PARTICIPANTS WITH SPEECH INTONATION PERCEPTION DEFICITS

To determine whether any amusics showed the same pattern of performance as reported by Lochy et al. (2004) for a subset of their amusics (namely, substantially *worse* performance on linguistic SQ discrimination than on tone sequence analogs), we examined the performance of each amusic on the two types of stimuli. As reported below, three amusics fit the profile identified by Lochy et al. Further data analysis focused on these three individuals.

MEASUREMENT OF FINAL PITCH FALL/RISE

To compare the acoustic characteristics of linguistic SQ items that did (vs. did not) cause difficulty for the selected amusics, we measured the final pitch fall (or rise) in all sentences used in this study. The frequencies (in Hz) of lowest and highest F0 values during the glide were measured for each sentence. Glide size (in st) was quantified from these values, and glide rate was computed by dividing glide size by the duration (in seconds) between these two pitch points (Figure 1). (Semitones

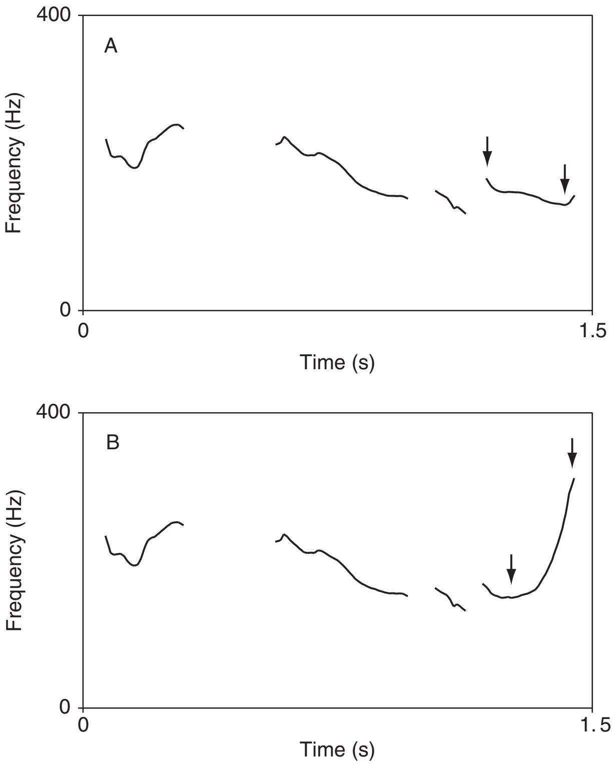

FIGURE 1. F0 contours of the sentence "He works ten hours a day" spoken as a statement (A) and as a question (B). Arrows mark the high and low points of the final pitch glides in the sentences.

were used as the unit of measurement of glide size in accordance with research on the perceptual scaling of intonation [Nolan, 2003]. If the lowest or highest F0 value of a glide was part of a pitch plateau, the onset/offset of the glide was defined as the boundary point between the plateau and the start of F0 change for the glide). Based on these data, mean glide size and glide rate were computed for trials that were classified correctly vs. incorrectly in the same-different discrimination task.

Results

GROUP DATA AND IDENTIFICATION OF AMUSICS WITH INTONATION PERCEPTION PROBLEMS

Table 2 shows performance of the British amusics on the SQ discrimination tasks. (Performance on the focus-shift tasks and composite MBEA scores are also shown for reference.)

For the SQ tasks the amusics did slightly *better* as a group on the tone sequences than on the speech stimuli, a surprising result given the well-known problems that amusics have with musical melody perception. However, this difference in performance was not significant (whether performance is measured as %Hits–%FA, or as % correct, Mann-Whitney *U prime* = 62, $n = 20$, $p = .36$ for both cases; all *p* values reported in the current paper were computed using the Mann-Whitney *U* test unless otherwise specified). Hence, one can conclude that the amusics as a group did about equally well on the linguistic and the nonlinguistic task. This finding contrasts with that of Ayotte et al. (2002), who found that amusics had no significant problem with linguistic SQ discrimination but had more difficulty discriminating tone sequence analogs of statements and questions. The current results replicate of the findings of Lochy et al. (2004) (cf. the Introduction and Study 2 below).

When performance on the SQ tasks was examined on an individual basis, three amusics showed a pattern resembling that shown by a subset of Lochy et al.'s (2004) amusics, namely substantially lower performance on speech vs. nonspeech discrimination. These three individuals scored 31% lower on speech than on tones on average, in terms of hits–false alarms (BA1-BA3, marked in bold in Table 2). When examined individually, their proportion of errors on sentences vs. tone sequences was not significantly different. However, when data from these three amusics were pooled, the proportion of correct vs. incorrect responses was significantly lower for speech than for tones ($\chi^2 = 10.3$, $n = 192$, $p < .01$, Fisher exact test, 2-tailed).

Notably, for BA1-BA3 the pattern of poorer performance on speech vs. nonspeech was not seen in focus-shift discrimination. Indeed, these amusics were relatively good at discriminating focus-shift sentences compared to their pronounced difficulty with linguistic SQ discrimination.

While the current study did not collect data from normal controls, Nicholson et al. (2003) tested 12 normal individuals (mean age = 70 years) on the same stimuli used in the current study and quantified performance as percent correct. Mean scores for SQ speech and tone discrimination were 94.8% and 93.5%, respectively (*SD* = 5.1%, 5.6%), suggesting that the amusics in the current study are discriminating SQ tone sequences

TABLE 2. Performance of British amusics (BA1-10) on Linguistic and Nonlinguistic Discrimination Tasks.

		SQ %H-%FA		SQ %Correct		FS % Correct		
Participant	Age	Speech	Tones	Speech	Tones	Speech	Tones	MBEA %Correct
BA1	**73**	**43.8**	**81.3**	**71.9**	**90.6**	**93.8**	**78.1**	**61.1**
BA2	**71**	**56.3**	**87.5**	**78.1**	**93.8**	**96.9**	**87.5**	**74.4**
BA3	**62**	**75.0**	**100**	**87.5**	**100**	**84.4**	**84.4**	**72.8**
BA4	30	75.0	87.5	87.5	93.8	93.8	87.5	70.1
BA5	58	75.0	75.0	87.5	87.5	75.0	75.0	63.3
BA6	60	81.3	93.8	90.6	90.6	100	81.3	71.7
BA7	55	93.8	100	96.9	100	100	75.0	67.8
BA8	65	93.8	75.0	96.9	87.5	90.6	81.3	58.3
BA9	51	100	81.3	100	90.6	93.8	90.6	53.3
BA10	58	100	93.8	100	96.9	100	93.8	73.9
Mean	58.3	79.4	87.5	89.7	93.8	92.8	83.4	66.7
SD	12.0	18.7	9.3	9.3	4.7	7.9	6.4	7.3

Note: Bold = substantially worse performance on speech than tones (on SQ items); Underlined = male; Italic = left handed; SQ = Statement-Question; FS = Focus-shift; %H - %FA = Percentage of hits – percentage of false alarms as well as normal controls. Nicholson et al.'s data also suggest that the pattern of substantially worse performance on linguistic vs. nonlinguistic SQ discrimination is not found in normal controls. In their study, 3 out of 12 control participants scored worse on speech than on tones, but by an average margin of just 6% (in terms of percent correct). Amusics BA1 – BA3 in the current study scored on average 16% worse on speech than on tones (in terms of percent correct), and thus seem to have an atypical degree of difficulty with linguistic vs. nonlinguistic SQ discrimination.

MEASUREMENTS OF GLIDE SIZE AND GLIDE RATE

Columns 3 and 4 of Table 1 show the sizes and rates of the final pitch glides in English statements and questions. Upward glides (questions) are about an octave in size on average, and are substantially larger in absolute size (*U prime* = 138, *n* = 24, *p* < .01) and faster in absolute rate (*U prime* = 135, *n* = 24, *p* < .01) than downward glides (statements).

RESULTS OF ACOUSTIC ANALYSES FOR CORRECTLY VS. INCORRECTLY DISCRIMINATED SENTENCES

Figure 2 shows the average size (in st) of the final pitch glide of statements and questions across all trials experienced by BA1-BA3, separated according to whether the trials were classified correctly or incorrectly. (Since all statements had falling pitch and all questions had rising pitch, absolute values are shown). As noted above, pitch glides for questions were substantially larger than those for statements on average, but the relevant finding for the current purposes is that glides in incorrect trials were not smaller than those in correct trials. In fact, there was a tendency for glides in the former trials to be larger in size, though this difference is not statistically significant for either statements (*U prime* = 595, *n* = 72, *p* = .24) or questions (*U prime* = 531, *n* = 72, *p* = .56).

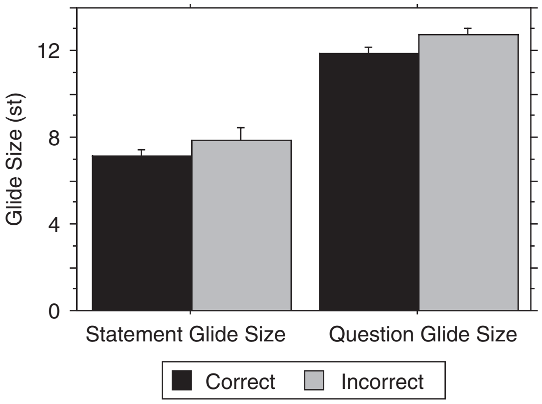

FIGURE 2. Glide sizes (mean and *SE* of absolute values, in st) for pitch movements at the ends of English statements and questions, separated according to trials classified correctly vs. incorrectly. Data from amusics BA1-BA3.

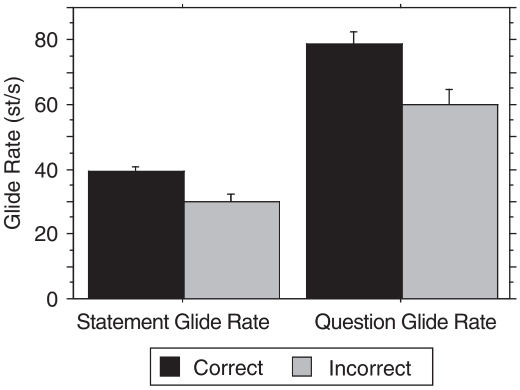

FIGURE 3. Glide rates (mean and *SE* of absolute values, in st/s) for pitch movements at the ends of English statements and questions, separated according to trials classified correctly vs. incorrectly. Data from amusics BA1-BA3.

In contrast, measurements of glide rate reveal that glides in incorrect trials were slower on average than those in correct trials (Figure 3). This difference is significant for both statements (*U prime* = 760, *n* = 72, *p* < .01) and questions (*U prime* = 691, *n* = 72, *p* < .01).

This difference in rate reflects a difference in the durations of glides in incorrect vs. correct trials: glides were significantly longer in incorrect than in correct trials for both statements (*U prime* = 739, *n* = 72, *p* < .01) and questions (*U prime* = 764, *n* = 72, *p* < .01). In correct trials, statement and question glides averaged 197 ms and 162 ms in duration, respectively (*SD* = 83, 54), while in incorrect trials these glides averaged 276 ms and 231 ms in duration, respectively (*SD* = 101, 72).

Further analysis of the errors committed by BA1-BA3 revealed that there were far more misses than false alarms (85% vs. 15% of the total errors, $\chi^2 = 12.4$, *n* = 96, *p* < .01, Fisher exact test, 2-tailed). Hence the primary problem was a failure to hear a difference between a linguistic statement and a question, rather than a tendency to perceive two identical statements (or questions) as different.

Study 2: French-Canadian Amusics

Method

The data for the French-Canadian amusics came from the original study of Lochy et al. (2004), as discussed in the Introduction. That study examined discrimination of focus-shift sentences in addition to statement-question

TABLE 3. Sentences Used in SQ discrimination with French-Canadian Participants.

Sentence	Rate (syl/s)	Size of Final Pitch Glide (st)		Rate of Final Pitch Glide (st/s)	
		S	Q	S	Q
Elle joue de la flute./?	6.1	−5.6	9.3	−20.2	29.4
Il veut partir maintenant./?	5.3	−4.8	9.4	−17.6	55.3
François est au restaurant./?	5.1	−3.7	10.9	−28.0	58.7
Le téléphone ne marche pas./?	5.6	−7.9	12.1	−49.5	93.6
Il travaille dix heures par jour./?	4.5	−3.6	11.4	−52.0	77.3
Elle a déjà lu ce livre./?	6.6	−4.0	17.0	−11.1	47.5
Il aime conduire des voitures rapides./?	6.1	−4.6	10.7	−17.5	44.6
Il était à Paris depuis trois mois./?	6.1	−5.6	14.3	−21.3	67.1
Le super-marché est fermé le dimanche./?	6.1	−4.5	15.3	−39.5	98.6
Il veut acheter une maison près de la plage./?	6.1	−4.1	11.2	−22.4	62.5
Mean	**5.8**	**−4.8**	**12.2**	**−27.9**	**63.4**
SD	0.6	1.3	2.6	14.2	21.6

Note: The syllable(s) on which pitch glided up (for questions) or down (for statements) is underlined in column 1.

sentences, as well as tone sequence analogs of both types of sentences.

PARTICIPANTS

Eleven French-Canadian amusics (7 women, mean age 60.1 years, *SD* = 6.7) and 11 matched controls free of neurological or psychiatric disorders participated in the study. All were right-handed and 7 of the amusics had participated in Ayotte et al.'s (2002) study. The amusics' musical deficit was confirmed using the MBEA (Peretz et al., 2003).

STIMULI

The materials for SQ discrimination consisted of 10 pairs of sentences uttered by a female native speaker of continental French (average F0 across all sentences = 244.3 Hz). Sentences ranged between 6 and 11 syllables (average duration = 1.4 s). Sentence pairs were created in the same manner as in British study.[3] Table 3 lists the sentences and gives some of their acoustic characteristics. For each sentence, a nonlinguistic tone sequence analog was created as described in Study 1 above.

PROCEDURE AND SCORING

From the original 10 pairs of sentences, a list of 40 pairs of test stimuli were created as follows: Each pair was presented in 'same' configuration in two forms (both statements or both questions, yielding 20 pairs), and in 'different' configuration in two forms (statement first, followed by question, and vice-versa, yielding 20 more pairs). These SQ pairs were mixed with other speech discrimination pairs in blocks of 30 pairs. A block only contained linguistic stimuli or tone analog stimuli. Blocks alternated between linguistic and tone stimuli, and order of block presentation was counterbalanced across participants. These blocks were presented for same-different discrimination.

Within each block members of a pair were separated by 750 ms and pairs were separated by 4.5 s. Participants were instructed to listen to each pair and indicate with a button press if the members of the pair sounded identical or if they sounded different in any way. Practice items were given before each block. Participants were tested in a quiet room. Performance is reported here in terms of hits and false alarms and in terms percentage of correct responses, for comparison with data from the British amusics.

IDENTIFYING PARTICIPANTS WITH SPEECH INTONATION PERCEPTION DEFICITS

To identify amusics with substantially worse performance on linguistic SQ discrimination than on tone sequence analogs, we examined the performance of each amusic on the two types of stimuli, as in Study 1.

---

[3] In two sentence pairs, there were minor differences in the average F0 of a single syllable preceding the critical region. In "Il aime conduire des voitures rapides./?" the syllable "ra" of "rapides" had an average F0 of 200 Hz in the statement vs. 210 Hz in the question. In "François est au restaurant./?" the syllable "res" of "restaurant" had an average F0 of 260 Hz in the statement vs. 242 Hz in the question. In each sentence pair, the shape of the F0 contour over the syllable in question was very similar, i.e., there was no difference in the direction of the pitch contour on this syllable.

MEASUREMENT OF FINAL PITCH FALL/RISE

To compare the acoustic characteristics of linguistic SQ items that did (vs. did not) cause difficulty for the selected amusics, we measured the final pitch fall (or rise) in all sentences used in this study, using the same procedures as in Study 1.

Results

GROUP DATA AND IDENTIFICATION OF AMUSICS WITH INTONATION PERCEPTION PROBLEMS

Table 4 shows performance of the amusics and controls on the SQ discrimination tasks. (Performance on the focus-shift tasks and composite MBEA scores for amusics are also shown for reference.)

As first noted by Lochy et al. (2004), for SQ discrimination the amusics as a group performed worse on sentences than on tone sequence analogs, a striking finding since Ayotte et al. (2002) reported the opposite pattern of results. It should be noted, however, that in the current study the difference is not statistically significant (whether performance is measured as %Hits - %FA or as % correct, *U prime* = 79, *n* = 22, *p* = .21 in both cases). Examination of the individual data reveals the reason for this lack of statistical significance: while performance on the tone sequence task was fairly uniform (and surprisingly good), performance on the linguistic task showed large individual variation. Four of the amusics had substantial difficulty with this task compared to the tone sequence task, scoring at least 30% worse on the sentences than on the tone sequences (data in bold in Table 4). In keeping with Study 1, further analyses of acoustic

TABLE 4. Performance of French-Canadian Amusics (CA1-11) and Controls (CC1-11) on Linguistic and Nonlinguistic Discrimination Tasks.

		SQ %H-%FA		SQ %Correct		FS %Correct		
Participant	Age	Speech	Tones	Speech	Tones	Speech	Tones	MBEA %Correct
Amusics								
CA1	**67**	**0.0**	**90.0**	**50.0**	**95.0**	**70.0**	**82.5**	**67.8**
CA2	**61**	**35.0**	**80.0**	**67.5**	**90.0**	**97.5**	**65.0**	**51.1**
CA3	**53**	**40.0**	**100**	**70.0**	**100**	**95.5**	**80.0**	**60.6**
CA4	**56**	**70.0**	**100**	**85.0**	**100**	**82.5**	**60.0**	**57.2**
CA5	67	75.0	95.0	87.5	97.5	—	—	58.9
CA6	64	85.0	100	92.5	100	—	—	70.0
CA7	45	95.0	100	97.5	100	92.5	92.5	66.1
CA8	64	100	100	100	100	100	100	68.9
CA9	64	100	80.0	100	90.0	90.0	57.5	57.8
CA10	62	100	80.0	100	90.0	92.5	92.5	71.7
CA11	58	100	95.0	100	97.5	—	—	73.3
Mean	60.1	72.7	92.7	86.4	96.4	90.1	78.8	63.9
SD	6.7	33.8	8.8	16.9	4.4	9.7	16.2	7.2
Controls								
CC1	64	85.0	95.0	92.5	97.5	—	—	
CC2	54	94.7	95.0	95.0	97.5	97.5	100	
CC3	45	95.0	90.0	97.5	95.0	—	—	
CC4	59	95.0	100	97.5	100	—	—	
CC5	68	95.0	90.0	97.5	95.0	100	100	
CC6	54	100	85.0	100	92.5	100	97.5	
CC7	59	100	90.0	100	95.0	100	97.5	
CC8	60	100	95.0	100	97.5	95.0	97.5	
CC9	56	100	95.0	100	97.5	—	—	
CC10	64	100	100	100	100	97.5	97.5	
CC11	52	100	100	100	100	97.5	97.5	
Mean	57.8	96.8	94.1	98.2	97.0	98.2	98.2	
SD	6.5	4.6	4.9	2.5	2.5	1.9	1.2	

Note: Bold = substantially worse performance on speech than tones (on SQ items); Underlined = male; SQ = Statement-Question; FS = Focus-shift; %H - %FA = Percentage of hits – percentage of false alarms; — = Missing data properties of sentences classified correctly vs. incorrectly are restricted to these amusics (though data from CA1 are excluded because he responded "same" to every pair of linguistic items, so that his data could not be used to compare acoustic characteristics of correctly vs. incorrectly discriminated SQ pairs). The remaining three amusics (CA2-CA4) scored 45% lower on average on sentence discrimination than on tone sequence discrimination (in terms of hits–false alarms). Each of these amusics had a significantly higher proportion of errors on linguistic SQ discrimination than on tone sequence discrimination (CA2: $\chi^2 = 6.1$, $n = 80$, $p < .03$; CA3: $\chi^2 = 20.1$, $n = 80$, $p < .001$; CA4: $\chi^2 = 9.5$, $n = 80$, $p < .03$, Fisher exact test, 2-tailed). This pattern of significantly worse performance on linguistic vs. nonlinguistic SQ discrimination was not found in any of the normal controls. Three out of 11 control participants scored worse on speech than on tones, but by an average margin of just 5% (in terms hits–false alarms).

Notably, CA2-CA4's pattern of poorer performance on speech vs. nonspeech was not seen in focus-shift discrimination Just as with BA1-BA3, CA2-CA4 were relatively good at discriminating focus-shift sentences, in contrast to their pronounced difficulty with linguistic SQ discrimination.

MEASUREMENTS OF GLIDE SIZE AND GLIDE RATE

Columns 3 and 4 of Table 3 show the sizes and rates of the final pitch glides in French statements and questions. As in Study 1 (English sentences), upward glides (questions) were about an octave in size on average, and were substantially larger in absolute size (*U prime* $= 100$, $n = 20$, $p < .01$) and faster in absolute rate (*U prime* $= 93$, $n = 20$, $p < .01$) than downward glides (statements).

RESULTS OF ACOUSTIC ANALYSES FOR CORRECTLY VS. INCORRECTLY DISCRIMINATED SENTENCES

Figure 4 shows the average size (in st) of the final pitch glide of statements and questions across all trials experienced by CA2-CA4, separated according to whether the trials were classified correctly or incorrectly by these amusics. (Since all statements had falling pitch and all questions had rising pitch, absolute values are shown). There was no statistically significant difference in glide size across incorrect vs. correct trials for either statements (*U prime* $= 919$, $n = 90$, $p = .97$) or questions (*U prime* $= 973$, $n = 80$, $p = .62$).

Figure 5 shows glide rate in incorrect vs. correct trials. In contrast to the finding with the British amusics, there was no significant difference in glide rate for incorrect vs. correct trials for either statements ($U = 919$, $n = 90$, $p = .97$) or questions (*U prime* $= 937$, $n = 90$, $p = .85$). Furthermore, glide duration also showed no systematic difference in the two kinds of trials. In correct trials, statement and question glides averaged 210 ms and 214 ms in duration, respectively ($SD = 89, 76$), while in incorrect trials these glides averaged 204 ms and 200 ms in duration, respectively ($SD = 81, 63$).

Further analysis of the errors committed by CA2-CA4 revealed that 100% of their errors were misses, with no false alarms. Hence, similar to the British amusics, their problem was a failure to hear a difference between a linguistic statement and a question rather than a tendency to perceive two identical statements (or questions) as different.

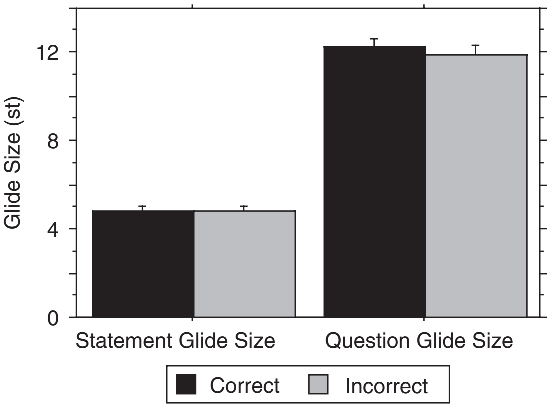

FIGURE 4. Glide sizes (mean and *SE* of absolute values, in st) for pitch movements at the ends of French statements and questions, separated according to trials classified correctly vs. incorrectly. Data from amusics CA2-CA4.

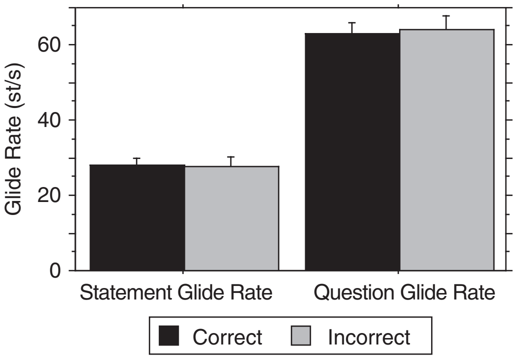

FIGURE 5. Glide rates (mean and *SE* of absolute values, in st/s) for pitch movements at the ends of French statements and questions, separated according to trials classified correctly vs. incorrectly. Data from amusics CA2-CA4.

Discussion

OVERVIEW: REEXAMINING SPEECH INTONATION PERCEPTION IN AMUSIA

While the first group study of prosody perception in musically tone-deaf individuals (Ayotte et al., 2002) suggested that amusics had no problems with the perception of speech intonation, the studies reported here have reexamined this issue. In two independent studies of British and French-Canadian amusics, about 30% of musically tone-deaf individuals have difficulty discriminating statements from questions on the basis of a final rising or falling pitch glide in an intonation contour. Yet strikingly (and paradoxically), these same individuals are able to discriminate sequences of discrete tones modeled on these intonation contours. These findings raise new questions about the domain specificity of amusia and about the nature of the underlying perceptual deficit(s) in this disorder.

From the standpoint of those interested in the domain specificity of music (e.g., Peretz & Coltheart, 2003), the current findings point to the need to incorporate tests of prosody perception into the diagnosis of amusia in order to separate those individuals with purely musical deficits from those with impairments spanning language and music. It is particularly important for genetic studies to be aware of possible subpopulations within amusia, some of whom may have nonmusical deficits (cf. Douglas & Bilkey, 2007).

The current findings raise numerous questions. First, why do the results differ from those of Ayotte et al. (2002), who found that amusics were generally good at linguistic SQ discrimination (and worse at the nonlinguistic analogs)? A difference between the procedures of Ayotte et al. and the current study of French-Canadian amusics is that the interval between members of a pair was shortened from 2 s to 750 ms, but it is not clear why this would lead to the current pattern of results. Second, why do only a subset of amusics (~30% of those tested here) exhibit intonation perception deficits? Does this subset have a qualitatively different kind of deficit from amusics without such problems, or are they simply at the end of a continuum such that all amusics would show intonation perception problems relative to controls if pitch contrasts were made subtle enough? Third, why do British but not French-Canadian amusics with intonation perception problems show sensitivity to pitch glide rate in terms of their pattern of discrimination errors?

In terms of the last question above, it should be noted that there were differences in the stimuli experienced by the two groups. For example, compared to the English sentences the French sentences were 38% faster in terms of syllables/second ($U\ prime = 452$, $n = 44$, $p < .01$), 30% higher in average F0 ($U\ prime = 477$, $n = 44$, $p < .001$), had 35% smaller final pitch falls in statements ($U\ prime = 106$, $n = 22$, $p < .01$), and were 26% slower on average in the rate of pitch falls in statements (marginally significant, $U\ prime = 89$, $n = 22$, $p < .06$) (cf. Tables 1 and 3). Furthermore, members of each stimulus pair were separated by only 2 s in the British study but only by 750 ms in the French-Canadian study. However, it is not clear how these differences can account for the fact that British amusics showed glide-rate dependency in their errors, while French-Canadian amusics did not.

To address these discrepancies, further work is needed using stimuli and tasks which are closely matched in acoustic properties across different groups of amusics. Furthermore, it would be desirable to manipulate glide size and glide rate in a parametric fashion to investigate the importance of these parameters to deficits in statement-question discrimination.

EVIDENCE FOR A PITCH DIRECTION DEFICIT IN SPEECH

A salient finding in the current work is that the amusics who had difficulty discriminating statements from questions did not have comparable difficulties discriminating between sentences with emphasis on different sentence-internal words ('focus-shift' sentences, cf. Tables 2 and 4). The linguistic focus-shift condition thus acts as an important partial control, because it suggests that that the amusics could detect pitch movements in speech. As long as these movements were on different words, amusics could use the movements to distinguish between sentences. However, when the movements were on the same word across the members of a sentence pair, but differed in direction (i.e., in the SQ sentences), these same amusics had difficulty.

This leads to the idea that amusics do in fact detect pitch movements in speech, but have difficulty perceiving the *direction* of these movements. A problem with the perception of pitch direction in speech is plausible given that psychophysical work has revealed a salient pitch direction determination deficit in amusia. For example, Foxton et al. (2004) had 13 British amusics and 10 controls judge the direction of a short (100 ms) pure tone pitch glide. Glide size was manipulated while glide duration was held constant (hence glide rate was also manipulated). The amusics had pronounced difficulty judging direction in smaller (hence slower) pitch glides: their average threshold for accurate direction judgments was 2.2 st, compared to 0.1 st for controls (corresponding to 22 st/s vs. 1 st/s, respectively).[4] It should be noted that the thresholds determined by Foxton et al. were based on pure tones centered on 500 Hz, and that pitch direction thresholds may depend on the frequency and spectral content of auditory stimuli. For the current purposes, the relevant point is that psychophysical work suggests that amusics have elevated thresholds for perceiving the direction of pitch glides, and that such a deficit appears to be independent of more basic deficits in simple pitch change detection (Griffiths, Stewart, McDonald, & Kumar, 2007; Patel, 2008; cf. Johnsrude, Penhune, & Zatorre, 2000). Such a deficit could be responsible for problems in linguistic SQ discrimination, where the critical cue is the direction of pitch change.

A RATE-BASED DIRECTION PITCH DEFICIT?

A novel aspect of the current findings is that the British amusics had difficulty discriminating statements from questions when the rate of change of the final pitch movement was relatively slow. Given that the French amusics did not show this pattern of results, more work is needed to test the robustness of this finding. As noted above, such work should control glide size and rate in a systematic fashion. For the moment, it is sufficient to note that a rate-based problem with pitch direction perception might explain why the British amusics examined in this study did well on discriminating tone sequence analogs of intonation. In each tone sequence pair, the critical difference is borne by the final pitch interval, in which pitch moves in a step-wise fashion between the final two tones. In step-like pitch changes the rate of pitch change is very fast, so that thresholds for pitch direction perception might be exceeded. (Alternatively, the use of step-like pitch changes may have made it possible for the British and French amusics to use the final pitch in the tone sequence analogs as a discrimination cue, since this pitch was always higher in question analogs than in statement analogs. Further work is needed to test these possibilities.)

Conclusion

The finding that some amusic individuals have intonation perception problems highlights the importance of studying the domain specificity of musical tone deafness. The results of the current studies suggest that some amusics have difficulty discriminating the direction of linguistically relevant pitch movements in speech. Further work should explore whether amusics are generally worse than controls at linguistic pitch direction discrimination, using stimuli in which glide size and/or glide rate are systematically varied. More generally, research on intonation perception in amusia provides a powerful way to probe the domain specificity of melodic processing in speech and music.

Author Note

Supported by Neurosciences Research Foundation as part of its research program on music and the brain at The Neurosciences Institute, where ADP is the Esther J. Burnham Senior Fellow. The research on French Canadian amusics was supported by grants from the Canadian Institutes of Health Research, the Human Frontier Science Program and the Natural Sciences and Engineering Research Council of Canada. We thank Lola Cuddy and an anonymous reviewer for helpful comments.

Correspondence concerning this article should be addressed to Aniruddh D. Patel, The Neurosciences Institute, 10640 John Jay Hopkins Dr., San Diego, CA 92121. E-MAIL: apatel@nsi.edu

---

[4] In computing these averages from the original data of Foxton et al. (2004), one amusic and one control were excluded from the analysis because their direction thresholds were > 2 *SD* larger than the rest of their group.

References

ALLEN, G. (1878). Note-deafness. *Mind*, *3*, 157-167.

AYOTTE, J., PERETZ, I., & HYDE, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238-251.

BOLINGER, D. (1985). *Intonation and its parts: Melody in spoken English.* London: Edward Arnold.

CUDDY, L. L., BALKWILL, L-L., PERETZ, I., & HOLDEN, R. R. (2005). A study of "tone deafness" among university students. *Annals of the New York Academy of Sciences, 1060,* 311-324.

DOUGLAS, K. M., & BILKEY, D. K. (2007). Amusia is associated with deficits in spatial processing. *Nature Neuroscience*, *10,* 915–921.

DRAYNA, D., MANICHAIKUL, A., DE LANGE, M., SNIEDER, H., & SPECTOR, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science*, *291,* 1969-72.

FOXTON, J. M., DEAN, J. L., GEE, R., PERETZ, I., & GRIFFITHS, T. D. (2004). Characterisation of deficits in pitch perception underlying 'tone deafness.' *Brain*, *127,* 801-810.

GÓSY, M., & TERKEN, J. (1994). Question marking in Hungarian: Timing and height of pitch peaks. *Journal of Phonetics, 22,* 269-281.

GRIFFITHS, T. D., STEWART, L., MCDONALD, C., & KUMAR, A. (2007, November) *Pitch perception and pitch memory in congenital amusia.* Poster session presented at the Society for Neuroscience Meeting, San Diego.

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. *Journal of Neuroscience, 27,* 13028-13032.

HYDE, K., & PERETZ, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15,* 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.

JOHNSRUDE, I. S., PENHUNE, V. B., & ZATORRE, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain, 123,* 155-63.

LADD, D. R. (1996). *Intonational phonology.* Cambridge: Cambridge University Press.

LOCHY, A., HYDE, K. L., PARISEL, S., VAN HYFTE, S., & PERETZ, I. (2004, April). *Discrimination of speech prosody in congenital amusia.* Poster session presented at the meeting of the Cognitive Neuroscience Society, San Francisco.

MANDELL, J., SCHULZE, K., & SCHLAUG, G. (2007). Congenital amusia: An auditory-feedback disorder? *Restorative Neurology and Neuroscience, 25,* 323-334.

NICHOLSON, K. G., BAUM, S., KILGOUR, A., KOH, C. K., MUNHALL, K. G., & CUDDY, L. L. (2003). Impaired processing of prosodic and musical patterns after right hemisphere damage. *Brain and Cognition, 52,* 382-389.

NIEBUHR, O. (2003). Perceptual study of timing variables in F0 peaks. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, 2003,* pp. 1225-1228.

NOLAN, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona* (pp. 771-774). Barcelona: Universitat Autònoma de Barcelona.

PATEL, A. D. (2008). *Music, language, and the brain.* New York: Oxford University Press.

PATEL, A. D., FOXTON, J. M., & GRIFFITHS, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition, 59,* 310-313.

PATEL, A. D., PERETZ, I., TRAMO, M. J., & LABRECQUE, R. (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language, 61,* 123-144.

PERETZ, I., AYOTTE, J., ZATORRE, R. J., MEHLER, J., AHAD, P., PENHUNE, V., & JUTRAS, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33,* 185-191

PERETZ, I., BRATTICO, E., & TERVANIEMI, M. (2005). Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology, 58,* 478-482.

PERETZ, I., CHAMPOD, 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., & COLTHEART, M. (2003). Modularity of music processing. *Nature Neuroscience, 6,* 688-691.

PERETZ, I., CUMMINGS, S., & DUBÉ, M-P. (2007). The genetics of congenital amusia (or tone-deafness): A family aggregation study. *American Journal of Human Genetics, 81,* 582-588.

SACKS, O. (2007). *Musciophilia: Tales of music and the brain.* New York: Knopf.