Journal of Experimental Psychology: General
2015, Vol. 144, No. 4, 730–736

0096-3445/15/$12.00 http://dx.doi.org/10.1037/xge0000081

BRIEF REPORT

Individual Differences in the Perception of Melodic Contours and Pitch-Accent Timing in Speech: Support for Domain-Generality of Pitch Processing

Tuuli H. Morrill
George Mason University

J. Devin McAuley, Laura C. Dilley, and David Z. Hambrick
Michigan State University

Do the same mechanisms underlie processing of music and language? Recent investigations of this question have yielded inconsistent results. Likely factors contributing to discrepant findings are use of small samples and failure to control for individual differences in cognitive ability. We investigated the relationship between music and speech prosody processing, while controlling for cognitive ability. Participants (n = 179) completed a battery of cognitive ability tests, the Montreal Battery of Evaluation of Amusia (MBEA) to assess music perception, and a prosody test of pitch peak timing discrimination (early, as in *insight* vs. late, *incite*). Structural equation modeling revealed that only music perception was a significant predictor of prosody test performance. Music perception accounted for 34.5% of variance on prosody test performance; cognitive abilities and music training added only about 8%. These results indicate musical pitch and temporal processing are highly predictive of pitch discrimination in speech processing, even after controlling for other possible predictors of this aspect of language processing.

Keywords: music, individual differences, cognitive ability, speech perception

The question of whether the same mechanisms underlie processing of music and language has long been of interest to psychologists, and is of particular interest to cognitive scientists (Patel, Wong, Foxton, Lochy, & Peretz, 2008; Zatorre, Belin, & Penhune, 2002). Evidence suggests that at least some aspects of music and language processing involve domain-general mechanisms (Koelsch et al., 2002; Maess, Koelsch, Gunter, & Friederici, 2001; Nicholson, Baum, Cuddy, & Munhall, 2002; Patel, Peretz, Tramo, & Labreque, 1998; Schön et al., 2010). Much of this work has focused on pitch processing, since pitch information comprises an integral part of melodic systems in music and prosodic systems in language. Evidence for similar perceptual abilities across domains comes from several sources. For example, tone language speakers have shown advantages in discriminating nonspeech pitch intervals (Giuliano, Pfordresher, Stanley, Narayana, & Wicha, 2011; Pfordresher & Brown, 2009; Wong et al., 2012). Conversely, musical training may affect linguistic pitch processing, in both tonal and nontonal languages (Bidelman, Gandour, & Krishnan, 2011; Magne, Schon, & Besson, 2006; Schön, Magne, & Besson, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007).

One influential line of work investigating domain generality of pitch has considered individuals diagnosed with congenital amusia. Congenital amusia is characterized by deficits in musical pitch processing that include difficulty discriminating small pitch changes, detecting the direction of a pitch change, and recognizing melodies without lyrics (Ayotte, Peretz, & Hyde, 2002; Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Hyde & Peretz, 2004). Less clear is whether individuals with amusia show impaired pitch processing in the domain of language. In speech, pitch and timing cues (i.e., prosodic cues) often convey word-level or sentence-level meaning. Thus, if pitch processing deficits are domain-general, rather than music-specific, prosodic processing would also be expected to be impaired in individuals with amusia.

Some studies investigating prosodic processing by individuals with amusia have found no differences from individuals without amusia (Ayotte et al., 2002; Peretz et al., 2002); these studies have generally focused on the broad-scale intonation patterns representing information structure and expressive meaning—for example, the falling or rising pitch at the ends of statements or yes-no questions in English. Others have revealed that individuals with amusia do show some impairment in prosodic processing, both in

---

Editor's Note. David Lubinski served as the action editor for this article.—IG

Tuuli H. Morrill, Department of English, George Mason University; J. Devin McAuley, Department of Psychology, Michigan State University; Laura C. Dilley, Department of Communicative Sciences and Disorder, Michigan State University; David Z. Hambrick, Department of Psychology, Michigan State University.

We thank Editor David Lubinski and three anonymous reviewers for comments and suggestions that greatly improved this manuscript. We would like to thank Alan Wedd, Elizabeth Wieland, Patrick Engler, Brad Seegert, and Erica Koelzer for their assistance with this project. We thank Louis Vinke for comments on an earlier version of this article. This research was supported in part by the Department of Psychology at Michigan State University.

Correspondence concerning this article should be addressed to Tuuli H. Morrill, Linguistics Program, Department of English, George Mason University, 4400 University Drive, 3E4, Fairfax, VA 22030. E-mail: tmorrill@gmu.edu

tone languages (with lexical tones) and in languages with phrasal intonation (Hutchins, Gosselin, & Peretz, 2010; Jiang, Hamm, Lim, Kirk, & Yang, 2010; Liu, Patel, Fourcin, & Stewart, 2010; Nan, Sun, & Peretz, 2010; B. Tillmann et al., 2011). Evidence from investigations of emotional prosody provides additional support for shared processing mechanisms (Lima & Castro, 2011; Thompson, Marin, & Stewart, 2012). While trained musicians exhibit increased sensitivity to emotional speech prosody (Lima & Castro, 2011), individuals with amusia have shown significantly decreased performance for detecting emotions in speech, especially those conveyed by changes in pitch (Thompson et al., 2012). Overall, the detection of fine-grained pitch changes (e.g., occurring within the duration of a syllable and smaller than one semitone) most consistently elicits poorer performance in individuals with amusia (Tillmann et al., 2011; Zatorre & Baum, 2012).

Here, we examined the relationship between pitch processing in music and speech, while statistically controlling for cognitive abilities. The prosody task manipulated pitch peak timing, which has not previously been tested in examinations of domain-general processing. Despite the characterization of pitch changes in speech as being larger and occurring at a slower rate than in music (Peretz & Hyde, 2003; Tillmann et al., 2011; Zatorre & Baum, 2012), spoken language perception also involves rapidly occurring pitch contour information. For example, the timing of a pitch peak on the first syllable of the word *insight* (noun) distinguishes it from *incite* (verb), with a pitch peak on the second syllable, usually a difference of less than a few hundred milliseconds—these differences in the timing of pitch peak and syllable alignment are integral to English prosody (e.g., Shattuck-Hufnagel, Ostendorf, & Ross, 1994; Silverman & Pierrehumbert, 1990; Watson, Tanenhaus, & Gunlogson, 2008). We addressed two main questions: Could detection of differences in pitch peak timing represent an area of domain-general pitch processing used in both music and language perception? If so, is the ability to detect pitch peak timing differences independent of cognitive ability?

We used an individual-differences approach in which a large sample of participants completed the Montreal Battery of Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003), a speech prosody test, and tests of fluid intelligence, crystallized intelligence, working memory capacity, and perceptual speed. Fluid intelligence refers to the ability to solve novel problems, and crystallized intelligence to acquired knowledge (Carroll, 1993). Working memory capacity is the ability to keep task-relevant information in an active state (Engle, 2002), and perceptual speed refers to the efficiency of extracting information from the external environment. We performed structural equation modeling (SEM) to determine whether music perception ability would positively predict performance in the speech prosody test, and if so, whether this relationship would remain significant even after controlling for the cognitive abilities. A critical advantage of using SEM to answer this question is that it permits one to test relationships at the level of latent variables, which capture variance common to multiple measures of intended constructs and are free of measurement error. Thus, conclusions can be drawn at the level of the theoretical constructs of interest.

Method

Participants. The participants were 179 native English speakers ($n = 179$, ages 18–49 years; $M = 19.9$, $SD = 3.8$) with self-reported normal hearing and varied levels of formal music training ($M = 3.7$ years, $SD = 3.8$). A typical sample size for a study of this type is 150–200; thus, we sought to test at least 150 participants, and ultimately tested 179. As described below, we tested a model with five factors predicting prosody perception. A post hoc power analysis revealed that a sample size of 179 provided adequate power (>.90) to detect a medium-size effect (.30) in this model.

Participants were recruited from the Psychology Department at Michigan State University and a community subject pool over the course of one year and received course credit or nominal payment. Participation was open to native speakers of English; data were collected during the entire period that the subject pool was available.

The sample was not representative of the general population, as nearly all of the participants (>90%) were enrolled in college. However, there was no indication of severe range restriction in cognitive ability in our sample. As examples, the means for operation span and symmetry span (tests of working memory capacity) were 56.7 ($SD = 14.5$) and 28.0 ($SD = 8.0$), compared with means of 53.1 ($SD = 14.9$) and 24.3 ($SD = 8.8$) for these tests in a large normative sample consisting of both college students and nonstudents (Redick et al., 2012); and the mean for Raven's progressive matrices was 9.2 ($SD = 3.2$) compared with a mean of 9.0 ($SD = 3.7$) in a larger sample ($N = 534$) consisting of both students and nonstudents (Unsworth et al., 2015); and the means for synonym vocabulary and antonym vocabulary were 4.1 ($SD = 2.2$) and 3.1 ($SD = 2.0$), respectively, compared with means of 5.1 ($SD = 2.7$) and 4.9 ($SD = 2.6$) for these tests in young adults from Salthouse and colleagues' longitudinal study of cognitive aging (Timothy A. Salthouse, personal communication, February 25, 2015).

Materials and Procedure

MBEA. The MBEA includes six subtests, with three subtests for melodic organization (Scale, Contour, and Interval), two subtests for temporal organization (Rhythm and Meter), and a single musical memory subtest (Memory). In the Scale, Contour, Interval, and Rhythm subtests, participants make same/different judgments about pairs of novel melodies: *Same* melodies are identical, whereas *different* melodies include one note which violates the relevant component (scale, contour, etc.). For the Meter subtest, participants identify the melody as a "march" or "waltz." For the Memory subtest, participants indicate whether a melody is *old* (heard earlier) or *new* (not heard).

Cognitive Ability Battery. All tests were computerized, except the paper-and-pencil tests of perceptual speed. Working memory capacity (Unsworth, Heitz, Schrock, & Engle, 2005): In *operation span,* the participant attempted to remember letters while solving arithmetic equations. In *symmetry span,* the participant attempted to remember the location of filled cells in a grid while making judgments about whether patterns were symmetrical.

Fluid intelligence. In *letter sets* (Ekstrom, French, Harman, & Dermen, 1976), each of 20 trials consisted of five sets of four letters (e.g., ABCD); the goal was to identify the set that was different from the others. The time limit was 5 min. In *matrix reasoning* (Raven, Raven, & Court, 1962), each of 18 trials consisted of a matrix in which all but one cell (lower-right) was filled with a pattern; the goal was to identify the pattern that fit in the missing cell. The time limit was 10 min.

Crystallized intelligence. In *general information* (Hambrick, Salthouse, & Meinz, 1999), each of the 30 items was a cultural knowledge question. In *vocabulary*, each item was a word; the goal was to pick the synonym (10 items) or antonym (10 items) from four options. The time limit was 5 min.

Perceptual speed (from Salthouse & Babcock, 1991). In *letter comparison,* each item was a pair of letter combinations. There were 24 items on each of two pages; the goal for each page was to complete as many items as possible in 30 s, identifying pairs as "same" or "different." In *pattern comparison,* each item was a pair of geometric patterns, with a horizontal line separating the patterns. There were 30 items on each of two pages; the procedure was identical to *letter comparison*.

Prosody test. The timing of a pitch peak was varied across the first two syllables within a nonsense word ('Lannamaraine') while maintaining the magnitude, direction and rate of the pitch excursion. Stimuli were adapted from Dilley and Heffner (2013); the F0 contour of the phrase was resynthesized using the PSOLA algorithm in Praat (Boersma & Weenink, 2012). A nonsense word was chosen so that a known word's prescribed pitch peak location would not affect perception. Participants were familiarized with two *anchor* versions in which a pitch peak occurred with early timing (i.e., during the syllable, *la-,* at 500 ms into the phrase) or late timing (i.e., during, *na-,* at 650 ms). Four *test* versions contained pitch peaks equally spaced in 30 ms intervals between 500 and 650 ms (see Figure 1).

Participants completed a 16-trial familiarization phase, hearing each anchor eight times, then a 64 trial test phase AXB task, judging whether X (the test item) was more like A (the first anchor) or B (the second anchor). As established for this test (Dilley & Heffner, 2013), listeners should exhibit a perceptual category boundary between 560 and 590 ms; test items with a peak earlier than the boundary are judged as more like A, and items with a peak after the boundary are judged as more like B. Higher accuracy (response according to category boundary) is expected for the anchor items (identical to A or B) than for test items.

Data Analysis

Subtests of the MBEA Considered in the analyses included Scale, Contour, Interval, Meter, and Rhythm.

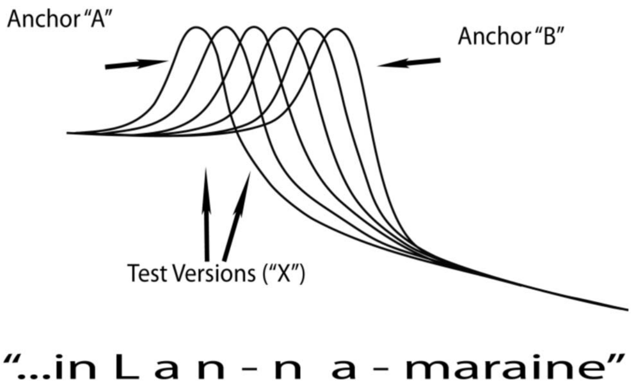

Figure 1. Schematic of pitch contour and peak alignment in anchor and test stimuli. Peaks were spaced at 30 ms intervals between 500 and 650 ms into the phrase (i.e., at 530, 560, 590, and 620 ms).

We screened the data for values more than 3.5 standard deviations (*SD*s) from the sample means. We replaced the 7 values that met this criterion with the 3.5 *SD* cutoff value, and used standardized (*z*) scores in the analyses. There were 5 missing values, which we replaced using the expectation maximization (EM) procedure in SPSS 19.0. All measures were approximately normally distributed (skewness, kurtosis values $< |2|$).

Performance on the prosody test was assessed separately for anchor items and test items. Average performance on the prosody test was 76% for the anchor items ($M = 75.9$ $SD = 16.7$) and 67% for the test items ($M = 67.5$, $SD = 13.2$). Listeners' average point of subjective equality (i.e., location of perceptual category boundary between early and late pitch peaks) was at 573 ms, approximately half-way between the early and late peak (575 ms), confirming the validity of assessing responses as accurate if they matched the anchor on the corresponding side of the category boundary.

Results

Descriptive statistics and correlations are in Table 1 and Table 2. The MBEA and cognitive ability measures correlated positively with each other, but the MBEA measures correlated more strongly with the prosody subscores (avg. $r = .37$) than did the cognitive ability measures (avg. $r = .19$). This pattern suggests that music perception predicted prosody perception, independent of cognitive ability. We used SEM to formally test this possibility.

Structural Equation Modeling

We report standard fit statistics for the SEM. The chi-square reflects the deviation between the observed and reproduced covariance matrices. The comparative fit index (CFI) and normed fit index (NFI) reflect improvement in model fit over a baseline model in which population covariances are assumed to be zero; values greater than .95 indicate good fit. The root-mean square error of approximation (RMSEA) reflects the difference between the observed correlations and the predicted correlations; values less than .08 indicate good fit (Kline, 2011).

We performed confirmatory factor analyses to establish a measurement model with latent variables representing the four cognitive ability factors—fluid intelligence (Gf), crystallized intelligence (Gc), working memory capacity (WMC), and perceptual speed (PS)—and two music perception factors—pitch processing and temporal processing. Model fit was good, $\chi^2(50) = 75.05$, $p = .012$, CFI = .96, NFI = .90, RMSEA = .05, but allowing the Gf and WMC indicators to load on a single factor did not significantly worsen model fit, $\Delta\chi^2(5) = 5.45$ ($p = .36$), nor did allowing the pitch processing and temporal processing indicators to load on a single factor, $\Delta\chi^2(5) = 5.19$ ($p = .39$). Therefore, we tested a model that included three cognitive ability factors—Gf/WMC, Gc, and PS—and a single music perception factor.[1] Model fit was good, $\chi^2(59) = 85.00$, $p = .015$,

---

[1] The question of whether Gf or WMC is more fundamental and causally prior remains open; therefore, to be theoretically neutral, we label the factor that includes both of these factors as Gf/WMC. We decided to exclude the memory subtest in the SEM, as at least two observed variables are required to model a latent factor and there was only this one memory subtest. Note, however, that the results of the SEM presented in Figure 2 are almost identical if we include the memory subtest as a sixth indicator of music perception (the effect of music perception on prosody perception increases from .61 to .64, *ps* $< .001$).

Table 1
Descriptive Statistics and Correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Fluid Intelligence																	
1. Letter sets	(.66)																
2. Matrix reasoning	.42	(.72)															
WM Capacity																	
3. Operation span	.35	.38	(.87)														
4. Symmetry span	.28	.35	.39	(.79)													
Crystallized Intelligence																	
5. General information	.27	.44	.25	.25	(.72)												
6. Vocabulary	.26	.33	.26	.25	.62	(.72)											
Perceptual Speed																	
7. Letter comparison	.20	.14	.14	.16	.21	.16	(.82)										
8. Pattern comparison	.39	.26	.22	.24	.26	.26	.42	(.72)									
Music perception																	
9. Contour	.15	.25	.28	.10	.20	.22	−.06	.08	(.66)								
10. Interval	.20	.24	.19	.08	.28	.28	−.01	.08	.69	(.67)							
11. Scale	.14	.22	.19	.08	.28	.30	−.03	.11	.60	.65	(.54)						
12. Rhythm	.28	.20	.28	.03	.24	.23	−.01	.15	.50	.63	.49	(.66)					
13. Meter	.22	.20	.08	.06	.32	.30	−.03	.13	.34	.23	.30	.28	(.88)				
14. Memory	.17	.27	.27	.18	.32	.27	.02	.11	.51	.49	.41	.53	.35	(.70)			
Prosody perception																	
15. Anchor	.18	.20	.32	.10	.25	.25	.18	.20	.44	.38	.31	.36	.24	.40	(.69)		
16. Test	.16	.16	.27	.05	.22	.19	.14	.22	.47	.44	.41	.39	.27	.35	.76	(.64)	
17. Yrs. of Music Training	.12	.18	−.01	.01	.21	.11	.04	.07	.16	.16	.29	.17	.36	.22	.10	.12	—
M	10.4	9.2	56.7	28.0	17.3	7.2	21.2	40.4	.78	.76	.82	.82	.82	.88	.76	.67	3.7
SD	2.9	3.2	14.5	8.0	4.4	3.7	5.4	6.2	.12	.13	.09	.12	.18	.10	.17	.13	3.8

Note. For descriptive statistics, correlations based on $N = 179$. WM = working memory; MP = music perception. Correlations with an absolute magnitude greater than .15 are statistically significant ($p < .05$). Values in parentheses along the diagonal are reliability estimates (coefficient alphas). The numbers along the top row of the correlation matrix correspond to the numbered tests at left.

CFI = .96, NFI = .89, RMSEA = .05. We accepted this model as the best-fitting measurement model (see Table 3 for latent variable correlations).

Next, we performed structural analyses to determine whether music perception positively predicted prosody perception. We tested for a direct effect of music perception on prosody perception, and then tested for this effect after statistically controlling for effects of the cognitive ability factors on both music perception and prosody perception. We also added a variable reflecting years of music training to the model to control for any influence of music training on the relationship between music perception and prosody perception. Results are displayed in Figure 2. As shown in Figure 2A, music perception had a strong positive effect on prosody perception (.59, $p < .001$), and as shown in Figure 2B, this effect was virtually unchanged, and even slightly higher (.61, $p < .001$), after controlling for effects of the other factors on music perception and prosody perception. music perception accounted for 34.5% of the variance in prosody perception, and the other predictors added only 7.9%. Overall fit of the model shown in Figure 2B was acceptable, $\chi^2(90) = 136.69$, $p = .001$, CFI = .95, NFI = .87, RMSEA = .05.

Discussion

We found that music perception is highly predictive of speech prosody perception—specifically, the processing of fine-grained prosodic cues—even after controlling for cognitive ability. This finding suggests that the ability to assess whether a pitch peak occurs relatively early or late in a contour may require a similar processing mechanism in both music and speech perception.

Table 2
MBEA Sub-Test Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Contour	179	.30	1.00	.78	.12
Interval	179	.43	1.00	.76	.13
Scale	179	.57	1.00	.82	.09
Rhythm	179	.37	1.00	.82	.12
Meter	179	.30	1.00	.82	.18
Memory	179	.47	1.00	.88	.10

Table 3
Latent Variable Correlations

	1	2	3	4	5
1. Fluid Intelligence/WMC	—				
2. Crystallized Intelligence	.63	—			
3. Perceptual Speed	.57	.43	—		
4. Music Perception	.40	.42	.11	—	
5. Prosody Perception	.33	.31	.32	.59	—

Note. WMC = Working Memory Capacity. All correlations are statistically significant ($p < .01$), except Perceptual Speed- Music Perception ($p = .29$).

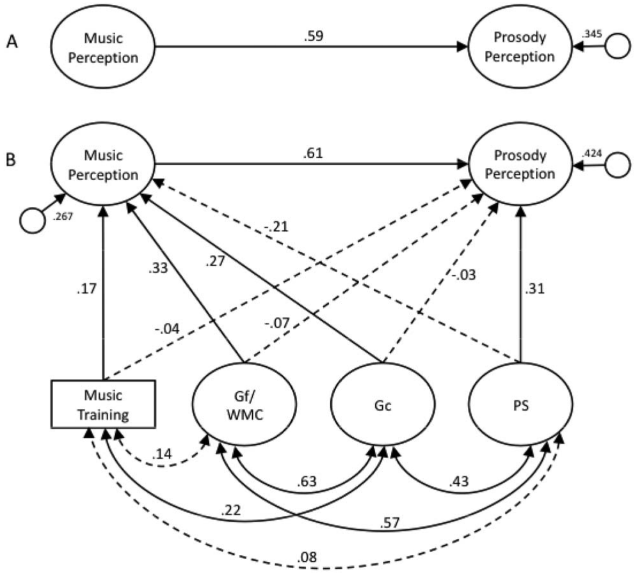

Figure 2. Structural model with music perception predicting prosody perception before (A) and after (B) controlling for cognitive ability factors and years of music training. Solid paths are statistically significant ($p <$ .05). Ovals represent latent variables. Values adjacent to doubled-headed arrows are correlation coefficients, values adjacent to single-headed errors are standardized regression coefficients (βs), and values adjacent to the small circles with arrows pointing to latent variables are squared multiple correlations (R^2s). The relationship between music perception and prosody perception is virtually unchanged after controlling for cognitive ability factors and music training ($\beta = .59$ vs. $\beta = .61$).

Unlike in many previous studies, this study did not exclusively examine individuals with music perception deficits, but instead used a sample of participants representing a wide range of cognitive and perceptual abilities. Another important aspect of this study pertains to the pitch-peak timing task, in which the fine-grained temporal and pitch manipulations were conducted on naturally produced (resynthesized) speech stimuli, whereas previous studies have generally used fully synthesized non-speech stimuli. Because there are discrepancies in the perception of pitch information in tonal versus speech stimuli even tonal stimuli that mimic the structure of speech, (Liu et al., 2012; Patel, Foxton, & Griffiths, 2005), using natural speech stimuli in conjunction with carefully controlled manipulations is important to interpretation of the results.

These results have implications for the nature of speech processing deficits in amusia. Individuals with deficits in music processing may have difficulty assessing the alignment of intonation patterns with speech structure and content (i.e., syllables, words, and phrases). Intonation patterns in speech are used to interpret meaning by associating pitch contours (e.g., pitch peaks), with specific locations in the speech stream. In a language like English, these contours can signal stressed and accented syllables that can distinguish word meaning; thus, perceptual difficulties associated with assessing pitch peak alignment may result in deficits in speech processing. Future research should include a broader assessment of speech prosody perception with multiple paradigms, to address potential limitations of a single measure (see, e.g., Gustafsson, 2002).

This finding will also be important for the identification of neural correlates associated with domain general processing mechanisms. It has been suggested that similar neural correlates of music and speech processing could be due to overlapping processes or task demands, such as the use of working memory and/or cognitive control (Kraus & Chandrasekaran, 2010; Rogalsky, Rong, Saberi, & Hickok, 2011; Zatorre & Baum, 2012). Our results provide evidence of music and speech prosody perception abilities which are correlated, but independent of other potentially overlapping task demands. The perceptual tasks used in the current study could be used to investigate neural correlates of processing mechanisms for music and speech perception.

Conclusion

The results of this study suggest that the ability to perceive changes in pitch-peak timing are controlled by a domain-general mechanism. In particular, the results suggest that similarities in the process of integrating fine-grained temporal and pitch information in music and speech are not accounted for by cognitive ability. These findings will be important for future investigations of neural correlates of domain general mechanisms.

References

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain: A Journal of Neurology, 125* (Pt. 2), 238–251. http://dx.doi.org/10.1093/brain/awf028

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23,* 425–434.

Boersma, P., & Weenink, D. (2012). *Praat: Doing phonetics by computer* [Computer program]. (Version 4.0.26). Retrieved from http://www.praat.org

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511571312

Dilley, L. C., & Heffner, C. (2013). The role of F0 alignment in distinguishing intonation categories: Evidence from American English. *Journal of Speech Sciences, 3,* 3–67.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19–23. http://dx.doi.org/10.1111/1467-8721.00160

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain: A Journal of Neurology, 127,* 801–810. http://dx.doi.org/10.1093/brain/awh105

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in Psychology, 2,* 146. http://dx.doi.org/10.3389/fpsyg.2011.00146

Gustafsson, J. (2002). Measurement from a hierarchical point of view. In H. Braun, D. Jackson, & D. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 73–95). New Jersey: Lawrence Erlbaum Associates.

Hambrick, D. Z., Salthouse, T. A., & Meinz, E. J. (1999). Predictors of crossword puzzle proficiency and moderators of age-cognition relations. *Journal of Experimental Psychology: General, 128,* 131–164. http://dx.doi.org/10.1037/0096-3445.128.2.131

Hutchins, S., Gosselin, N., & Peretz, I. (2010). Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Frontiers in Psychology, 1,* 236. http://dx.doi.org/10.3389/fpsyg.2010.00236

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15,* 356–360. http://dx.doi.org/10.1111/j.0956-7976.2004.00683.x

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia, 48,* 2630–2639. http://dx.doi.org/10.1016/j.neuropsychologia.2010.05.009

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *NeuroImage, 17,* 956–966. http://dx.doi.org/10.1006/nimg.2002.1154

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11,* 599–605. http://dx.doi.org/10.1038/nrn2882

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11,* 1021–1031. http://dx.doi.org/10.1037/a0024521

Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers. *PLoS ONE, 7*(2), e30374. http://dx.doi.org/10.1371/journal.pone.0030374

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain: A Journal of Neurology, 133* (Pt. 6), 1682–1693. http://dx.doi.org/10.1093/brain/awq089

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4,* 540–545.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18,* 199–211. http://dx.doi.org/10.1162/jocn.2006.18.2.199

Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain: A Journal of Neurology, 133,* 2635–2642. http://dx.doi.org/10.1093/brain/awq178

Nicholson, K. G., Baum, S., Cuddy, L. L., & Munhall, K. G. (2002). A case of impaired auditory and visual speech prosody perception after right hemisphere damage. *Neurocase, 8,* 314–322. http://dx.doi.org/10.1076/neur.8.3.314.16195

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. http://dx.doi.org/10.1016/j.bandc.2004.10.003

Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language, 61,* 123–144. http://dx.doi.org/10.1006/brln.1997.1862

Patel, A. D., Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception, 25,* 357–368. http://dx.doi.org/10.1525/mp.2008.25.4.357

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33,* 185–191. http://dx.doi.org/10.1016/S0896-6273(01)00580-3

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999,* 58–75. http://dx.doi.org/10.1196/annals.1284.006

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences, 7,* 362–367. http://dx.doi.org/10.1016/S1364-6613(03)00150-5

Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception & Psychophysics, 71,* 1385–1398. http://dx.doi.org/10.3758/APP.71.6.1385

Raven, J. C., Raven, J., & Court, J. (1962). *Coloured progressive matrices*. Oxford, UK: Oxford Psychologists Press.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28,* 164–171. http://dx.doi.org/10.1027/1015-5759/a000123

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *The Journal of Neuroscience, 31,* 3843–3852. http://dx.doi.org/10.1523/JNEUROSCI.4515-10.2011

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763–776. http://dx.doi.org/10.1037/0012-1649.27.5.763

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J.-L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage, 51,* 450–461. http://dx.doi.org/10.1016/j.neuroimage.2010.02.023

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41,* 341–349. http://dx.doi.org/10.1111/1469-8986.00172.x

Shattuck-Hufnagel, S., Ostendorf, M., & Ross, K. (1994). Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics, 22,* 357–388.

Silverman, K., & Pierrehumbert, J. (1990). The timing of prenuclear high accents in English. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 71–106). New York: Cambridge University Press.

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences, USA, 109,* 19027–19032. http://dx.doi.org/10.1073/pnas.1210344109

Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., & Peretz, I. (2011). Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages. *Frontiers in Psychology, 2,* 120. http://dx.doi.org/10.3389/fpsyg.2011.00120

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498–505. http://dx.doi.org/10.3758/BF03192720

Unsworth, N., Redick, T. S., McMillan, B. D., Hambrick, D. Z., Kane, M. J., & Engle, R. W. (2015). Is playing video games related to cognitive abilities? *Psychological Science*. Advance online publication. http://dx.doi.org/10.1177/0956797615570367

Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in on-line comprehension: H^* vs. $L+H^*$. *Cognitive Science, 32,* 1232–1244. http://dx.doi.org/10.1080/03640210802138755

Wong, P. C., Ciocca, V., Chan, A. H., Ha, L. Y., Tan, L.-H., & Peretz, I. (2012). Effects of culture on musical pitch perception. *PLoS ONE, 7* (4), e33424. http://dx.doi.org/10.1371/journal.pone.0033424

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10,* 420–422.

Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology, 10*(7), e1001372. http://dx.doi.org/10.1371/journal.pbio.1001372

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6,* 37–46. http://dx.doi.org/10.1016/S1364-6613(00)01816-7

Received October 1, 2014
Revision received March 6, 2015
Accepted April 10, 2015 ■