Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Musicality and non-native speech sound processing are linked through temporal, pitch and spectral acuity

Permalink

https://escholarship.org/uc/item/0933w4qt

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN

1069-7977

Authors

Kempe, Vera
Bublitz, Dennis
Brooks, Patricia J.

Publication Date

2013

Peer reviewed

Musicality and non-native speech sound processing are linked through temporal, pitch and spectral acuity.

Vera Kempe (vkempe@abertay.ac.uk)
Division of Psychology, School of Social and Health Sciences, Bell Street
Dundee, DD1 1HG, United Kingdom

Dennis Bublitz (dennis.bublitz@gmail.com)
Department of Psychology, College of Staten Island, City University of New York
Staten Island, NY 10314 USA

Patricia J. Brooks (patricia.brooks@csi.cuny.edu)
Department of Psychology, College of Staten Island, City University of New York
Staten Island, NY 10314 USA

Abstract

Are observed links between musicality and non-native speech sound processing due to superior sensory processing of temporal, pitch, and spectral information, which benefits both musical and linguistic processing? Native English speakers discriminated Norwegian tonal contrasts, non-linguistic pure-tone analogues, Norwegian vowels, and short tones differing in temporal, pitch and spectral characteristics. Musicality was measured using Gordon's (1989) Advanced Measures of Musical Audiation (AMMA). After controlling for effects of sex, non-verbal IQ and previous language learning experience, the link between AMMA scores and tonal contour discrimination was partially mediated by pitch acuity. In addition, tonal contrast, pitch contour and vowel discrimination were predicted by temporal and spectral acuity. No independent effects of musical training were found. Thus, links between musicality and non-native speech sound processing appear to be mainly due to superior temporal, pitch or spectral acuity, which, in turn, may play somewhat different roles in processing different speech sounds.

Keywords: Non-native phoneme processing; temporal acuity, pitch acuity; spectral acuity; musicality; tonal contrast; vowel contrast.

Introduction

A number of studies have documented links between musicality and the ability to discriminate non-native speech sounds (Delogu, Lampis, Belardinelli, 2006, 2010; Marie, Delogu, Lampis, Belardinelli & Besson, 2011; Slevc & Miyake 2006; Wong, Skoe, Russo, Dees & Kraus, 2007). These studies typically use complex psychometric measures of musical aptitude, which leaves open the question as to which specific sub-components of musical aptitude are associated with non-native speech sound processing. Tests like the Wing test (Wing, 1968) or Gordon's Advanced Measures of Musical Audiation (AMMA) (Gordon, 1989) rely on working memory for musical and rhythmic phrases as well as on the ability to discriminate subtle differences in pitch, timbre, intensity, and rhythm. It is not clear which of these sub-components of musical aptitude are linked to non-native sound processing, especially because different types of acoustic information may be important for distinguishing different types of speech sounds. Specifically, the perception of vowels, which differ in spectral information associated with the first and second formants, should be most strongly predicted by sensitivity to timbre. In contrast, the perception of consonants, which are often distinguished by temporal information such as Voice Onset Time or formant transitions, should benefit from sensitivity to rapid temporal changes. Finally, lexical tones require sensitivity to pitch and, to the extent that they encompass differences in pitch contour, also sensitivity to temporal information. Thus, different aspects of auditory sensory acuity may be important for the processing of different types of non-native speech sounds. So far, the relationship between musicality and non-native speech sounds processing has been quite consistently established for tonal contrasts (Marie, Delogu, Lampis, Belardinelli & Besson, 2011; Slevc & Miyake, 2006; Wong, Skoe, Russo, Dees & Kraus, 2007), but the findings are less clear for phonological contrasts (Delogu, Lampis & Belardinelli, 2006, 2010).

A primary aim of this study was therefore to examine the specific contributions of temporal, pitch, and spectral acuity to the processing of different non-native speech sounds, and to determine whether general measures of musical aptitude can explain additional variance in non-native speech sound processing above what is explained by these basic sensory processes. To this end, we examined both a tonal and a vowel contrast that exist in Norwegian, a language unfamiliar to our participants. Many dialects of Norwegian have lexical tone such that rising and falling-rising pitch accents distinguish minimal pairs of segmentally identical bi-syllabic words. For example, 'Hammer', spoken with the rising tone, is a Norwegian proper noun while 'hammer', spoken with the falling-rising tone, denotes the tool. These contrasts encompass temporal changes in fundamental frequency in the range of several hundreds of milliseconds. Norwegian also contains a vowel contrast not present in English, the /i/ - /y/ contrast. The existence of these Norwegian contrasts offers the possibility to use linguistic stimuli rather than isolated synthesized segments, which may be processed in ways that differ from processing of natural linguistic materials. Moreover, to control for differences between linguistic and non-linguistic stimuli, we also used the extracted pure-tone analogues of the Norwegian tonal contrasts as stimuli.

Musicality is a complex construct encompassing musical aptitude as well as musical expertise (Nardo & Reiterer, 2009). Studies comparing non-native speech sound processing between musicians and non-musicians (Marie et al., 2011; Wong et al., 2007) suggest that musical practice hones abilities such as sensory acuity, working memory, or attentional control, which may transfer to the linguistic domain. However, it is also possible that certain sensory or cognitive abilities benefit both musical and linguistic processing. To see whether musical expertise incurs benefits for non-native speech sound processing in addition to benefits associated with superior auditory sensory acuity, we also examined whether the duration of musical training would explain variance in non-native speech sound processing over and above measures of musical aptitude.

Method

Native speakers of English completed AX discrimination tasks for non-native tonal and vowel contrasts, as well as for synthesized sounds differing in temporal, pitch, and spectral characteristics. Musical aptitude was tested using Gordon's AMMA (Gordon, 1989). To control for non-verbal intelligence we administered Cattell's Culture Fair Intelligence Test (Cattell & Cattell, 1973). Music and language background questionnaires inquired about length of musical training and number of languages learned at home or at school, and elicited self-ratings of proficiency in each language (L2 and L3).

Participants

One hundred and three native speakers of American English (58 women, 45 men) aged 19-22 years participated in the study. Three participants failed to provide L3 proficiency self-ratings, and one participant failed to provide pitch discrimination data. These participants are missing from analyses including these variables.

Materials and Tasks

Advanced Measures of Musical Audiation (AMMA): Gordon's (1989) AMMA consists of 30 items, each of which comprises a short musical 'statement' followed after four seconds by a short 'answer' of the same length. These items contain either one or more tonal changes, or one or more rhythmic changes, but not both. Participants have to decide whether the phrases are the same or different. For 'different' items, participants are asked to decide whether the difference is a tonal or rhythm change. The test yields a tonal and rhythm score, as well as a composite score.

AX ('same-different') – tasks: All AX tasks comprised 32 'same' and 32 'different' trials. For the Norwegian tonal and vowel contrasts, 'same' trials comprised different within-category instantiations obtained from repeated recordings of the same word.

Temporal Acuity. We synthesized eight 250 Hz sinusoidal carrier waves with an overall duration of 600 ms differing in amplitude envelope onset rise times, and otherwise devoid of segmental, spectral and pitch information. The onset of the amplitude envelope was faded in with rise times to reach maximum amplitude at 0 ms, 10 ms, 20 ms, 30 ms, 60 ms, 70 ms, 80 ms and 90 ms. 'Different' trials comprised pairs of sounds differing in rise times by 60 ms (e.g. 0 ms vs. 60 ms or 10 ms vs. 70 ms etc.), centered around 45 ms, a value which has been reported as the category boundary between 'bowed' and 'plucked' sounds (Cutting & Rosner, 1974).

Pitch Acuity. We created eight 500 ms pure tone sinusoidal carrier waves ranging from 100 to 3000 Hz in steps increasing by 100 Hz resulting in tones of 100 Hz, 200 Hz, 400 Hz, 700 Hz, 1100 Hz, 1600 Hz, 2200 Hz, and 3000 Hz, as well as contrasts with a frequency increased by 2% resulting in tones of 102 Hz, 204 Hz, 408 Hz, 714 Hz, 1122 Hz, 2244 Hz, and 3060 Hz. The cumulative increase was designed to create sound pairs that subjectively sampled the pitch range at roughly similar intervals, taking into account the non-linearity of pitch perception. For the 'different' trials, each sound was paired with its corresponding contrast sound resulting in pairs of 100 Hz vs. 102 Hz, 200 Hz vs. 204 Hz, 400 Hz vs. 408 Hz etc.)

Spectral Acuity. To test spectral acuity, we incorporated the pure tones created for the pitch acuity test into complex tones comprising low (e.g. 100 Hz or 200 Hz), middle (e.g. 700 Hz or 1100 Hz) and high (e.g. 2200 Hz or 3000 Hz) frequencies. These frequencies were chosen to broadly mimic the fundamental frequency and the first two formants of speech, which are crucial for vowel perception. For 'different' pairs, one of the component tones was increased by 2%, and this change affected either the middle or the high frequency. For example, a 'different' pair might include a complex tone consisting of frequencies of 100 Hz, 1100 Hz and 3000 Hz and a complex tone consisting of frequencies of 100 Hz, 1122 Hz and 3000 Hz.

Norwegian tonal contrast. Recordings by a male native speaker of Norwegian of eight minimal pairs of Norwegian words containing a contrast between rising and falling-rising tonal contours were taken from Kempe, Thoresen, Kirk, Schaeffler & Brooks (2012)**.** Four pairs contained short vowels in the first (stressed) syllable (mean length 64 ms); the remaining four pairs contained long vowels (mean length 187 ms). Crucially, words with rising and with falling-rising tones did not differ in length of the first vowel (118 vs. 133 ms, p = .5), overall word length (396 vs. 417 ms, p = .2), and metric stress; thus, duration and stress could not be used as additional cues. Corresponding short and long vowel pairs were matched for their initial phoneme.

To ensure that a male advantage in the discrimination these tonal contrasts, as found in previous research (Kempe et al., 2012), was not an artifact of the male voice presenting the stimuli, we also created a female voice version of the stimuli. To control for indexical features, we submitted the male voice stimuli to the voice gender change algorithm in PRAAT (Boersma & Weenink, 2011) using a fundamental frequency of 220 Hz and scaling the first formant up by 20 %. All results below are averaged over the male-voiced and the female-voiced version of the AX-task.

Pitch contour (non-speech analogue of tonal contrast). The non-speech equivalents of the Norwegian tonal contrast comprised sine waves with pitch contours extracted from the fundamental frequency modulation of both the male-voiced and the female-voice Norwegian tonal contrasts. These stimuli contained no information other than the pitch contour of the Norwegian tones. Again, results below are averaged over the male-voiced and the female-voiced version of the AX-task.

Norwegian vowel contrast. We used eight minimal pairs of Norwegian mono-syllabic words containing the vowel /i:/ or /I/ vs. /y:/ or /Y/, a contrast between high front unrounded and rounded vowels which does not exist in English. Recordings of a male native speaker of Norwegian were taken from Kempe et al. (2012). Four word pairs contained the short vowels /I/ and /Y/ (mean length 67 ms), and the remaining four word pairs contained the long vowels /i:/ and /y:/ (mean length 150 ms). On average both members of a minimal pair did not differ in vowel length (108 vs. 108 ms, p = .9); thus, duration could not serve as additional cue.

Other measures: Participants also completed Cattell's Culture-Fair Test of Nonverbal Intelligence, Scale 3, Form A (Cattell & Cattell, 1973), a music background questionnaire on which they provided information about duration of musical training, and a language background questionnaire on which they indicated the number of languages learned, and rated their reading, writing, speaking and comprehension abilities in all languages on a scale from 1 (very poor) to 6 (native-like).

Procedure

AX discrimination tasks were presented in three blocks, with the first block containing the temporal, pitch and spectral AX tasks, the second block containing the male-voiced and female-voiced tonal AX tasks as well as the vowel AX task, and the third block containing the two AX tasks presenting the extracted pitch contours of the male-voiced and female-voiced Norwegian tonal stimuli. The fixed block sequence ensured that variance due to order effects was not confounded with participant variance, although task order was counterbalanced within blocks. AMMA and Culture Fair Intelligence Test were interspersed between blocks with their order counterbalanced as well. Informed consent was obtained and background questionnaires were completed prior to any of the tasks.

In each of the AX-tasks, participants received eight practice trials with feedback, followed by 64 test trials without feedback. Within a trial, sound stimuli were separated by an inter-stimulus interval of 200 ms; the inter-trial interval was 500 ms. Participants were asked to press the 's' key if they perceived the sounds to be the same and the 'd' key if they perceived them to be different.

Results

Participants' performance on the AX-tasks was converted into A', a sensitivity measure that corrects for differences in response bias, and ranges from 0 to 1, with 0.5 corresponding to chance. Table 1 shows performance with tonal contrasts, non-speech contour analogues, and vowels. As previous research had shown a male advantage for the processing of some non-native speech sounds (Kempe et al., 2012; Bowles, Silbert, Jackson & Doughy, 2011), results are given for male and female participants separately. A 3 (Condition) x 2 (Sex) ANOVA yielded a main effect of Condition, F(2,202) = 8.2, p < .001. Bonferroni-corrected post-hoc tests indicated that performance was superior for the vowels compared to both tonal conditions, all t's > 3.4, all p's < .01, which did not differ from each other, p = .7. The main effect of Sex, F(1,101) = 2.3, p = .1, and the interaction, F(2,202) = 2.8, p = .06, fell short of significance. This trend towards a male advantage in processing tonal contours, but not vowels confirms the previous findings (Kempe et al., 2012).

Table 1: Mean A' and standard deviations (in parentheses) for discrimination of Norwegian tonal contrasts, extracted pitch contours and Norwegian vowels.

	tonal contrast	pitch contour	vowel contrast
males	0.781 (0.103)	0.790 (0.087)	0.800 (0.101)
females	0.748 (0.097)	0.747 (0.104)	0.802 (0.102)

Zero-order correlations between the predictors of non-native speech-sound processing are provided in Table 2. Noteworthy findings involve a positive correlation between non-verbal intelligence and both AMMA scores, indicating that comparison of musical phrases relies to some extent on mechanisms shared with psychometric intelligence, such as working memory and cognitive control (Duncan, Emslie, Williams, Johnson & Freer, 1996). Also, as expected, pitch and spectral acuity were positively correlated with both AMMA scores. In contrast, temporal acuity was not correlated with the AMMA scores, which may reflect the fact that the temporal processing relevant for music involves a longer time scale than the rapid temporal changes in the order of tens of milliseconds presented in our temporal acuity test. Instead, temporal acuity was positively correlated with non-verbal intelligence, confirming the documented link between rapid temporal auditory processing and psychometric intelligence (Rammsayer & Brandler, 2007).

The results of multiple regression analyses of non-verbal intelligence, language background, sex (dummy-coded) and the musicality measures on performance with tonal contrasts, pitch contours and vowels are presented in Table 3 (upper panel). We found that non-verbal intelligence showed a trend towards a positive association with discrimination of tonal contours and non-linguistic pitch contours. As indicated above, there was also a statistically marginal male advantage for these stimuli. Crucially, the

AMMA rhythm score predicted performance with tonal contrasts, even if they were extracted and presented without linguistic information. For non-native vowels, only self-rated proficiency in L3 predicted performance. Thus, while tonal performance was related to musicality, vowel performance was not, confirming observations by Delogu et al. (2006; 2010). Note that there was no effect of musical training.

Next, we added temporal, pitch and spectral acuity to the model (lower panel of Table 3) to see whether auditory acuity explains the link between musicality and non-native tonal processing. All Variance Inflation Factors were below 3.8, suggesting that multi-collinearity was not a problem in this data set. For the tonal contrast, we found a significant effect of pitch acuity; the effects of temporal and spectral acuity fell short of significance. For the extracted pitch contour, we found a significant effect of spectral acuity; the effect of temporal acuity fell short of significance. For the vowel contrast, we found significant effects of temporal and spectral acuity. In other words, the data showed a tendency for temporal and spectral acuity to predict discrimination of all non-native speech contrasts while the predictive effect of pitch acuity was confined to the tonal contrasts.

Table 2: Zero-order Pearson correlations between predictor variables. CFI – Culture Fair Intelligence test, # of Ls – number of learned languages, N's range from 99 and 103 depending on missing values, ** p < .01, * p < .05.

	2	3	4	5
1. CFI	.25*	.03	.23*	.09
2. # of Ls		.42**	.79**	.04
3. L2 (rating)			.56**	.17
4. L3 (rating)				.12

	6	7	8	9	10
1. CFI	.22*	.30**	.24*	.14.	.07
2. # of Ls	.11	.19	.24*	.11	.12
3. L2 (rating)	.06	.03	.13	.06	.07
4. L3 (rating)	.19	.14	.21*	.11	.11
5. music (years)	.28**	.25*	.13	.09	.28**
6. tonal score		.71**	.19	.23*	.25*
7. rhythm score			.14	.34**	.32**
8. temporal				.14	.12
9. pitch					.50**
10. spectral					

To test explicitly whether the association of the AMMA rhythm score with the tonal and the extracted pitch contour contrasts was mediated by auditory acuity, we performed mediation analyses employing bootstrapping to estimate the 95% confidence intervals of the indirect effect using a procedure introduced by Hayes and Preacher (2013) for multiple predictor variables (SPSS-macro MEDIATE, http://www.afhayes.com/spss-sas-and-mplus-macros-and-code.html). A relative indirect effect is deemed to be statistically significant at p = .05 if these confidence intervals do not include zero. This analysis revealed that the effect of the AMMA rhythm score on the tonal contrast was partially mediated by pitch acuity (the obtained lower and upper boundaries of the confidence interval were .0001 and .0048, respectively). For the extracted pitch contour, there were no indirect effects.

Table 3: Standardized regression coefficients and proportion of variance accounted for in regression analyses with performance on tonal contrasts, pitch contours and vowel contrasts as criterion variables and sex, non-verbal intelligence, language background and musical ability measures as predictors at the first step (upper panel) and temporal, pitch and spectral acuity added at the next step (lower panel), ***p < .001, **p < .01, * p < .05, †p < .1.

	tonal contrast	pitch contour	vowel contrast
sex	-.16†	-.19†	.02
CFI	.16†	21*	.06
# of Ls	-.04	-.06	-.17
L2 (rating)	.15	-.07	.03
L3 (rating)	.03	.06	.36*
music (years)	.09	.00	-.01
tonal score	-.03	.01	.16
rhythm score	.40**	.30*	.11
adj. R^2	.21	.15	.09
F(9,97)	4.29***	3.10**	2.23*

	tonal contrast	pitch contour	vowel contrast
sex	-.10	-.18†	.02
CFI	.16†	.19†	.04
# of Ls	-.06	-.10	-.23
L2 (rating)	.16	-.06	.04
L3 (rating)	.00	.05	.36*
music (years)	.02	-.07	-.10
tonal score	-.09	-.03	.09
rhythm score	.32*	.26†	.10
temporal	.17†	.18†	.29**
pitch	.22*	.03	-.05
spectral	.19†	.22*	.26*
adj. R^2	.34	.20	.20
F(9,97)	5.54***	3.23**	3.23**

Discussion

When different measures of auditory sensory acuity relevant to non-native speech-sound processing were added into a multiple regression model, they had an independent effect beyond effects of musical aptitude. For Norwegian tonal contrasts, the effect of musical aptitude was partially mediated by pitch acuity. For non-linguistic pitch contours and for vowels, performance was mainly predicted by temporal and spectral acuity rather than musical aptitude. This suggests that associations between musical aptitude and non-native speech-sound processing predominantly arise from superior sensory processing encompassing the ability to make subtle distinctions in temporal, pitch, and spectral properties of the sounds. This finding is incompatible with claims that musical and linguistic processing exploits different cues—with language mainly relying on rapid temporal processing and music relying on processing of pitch and spectral information (Zatorre, Belin & Penhune, 2002). By showing that musical aptitude contributes little to non-native speech sound processing beyond effects of temporal, pitch and spectral acuity, our findings underscore the importance of these basic sensory processes for both music and speech sound processing, and thereby support the idea of partially shared mechanisms (Patel, 2003; Strait, Hornickel & Kraus, 2011). This is not to say that musical aptitude does not encompass other components that may or may not be shared with language; rather, we suggest that the links between musical aptitude tests and non-native speech sound processing reported in the literature may be due to basic temporal, pitch and spectral acuity to a significant degree.

Before sensory acuity measures were added to the regression model, effects of musical ability were found only for tonal contrasts and their non-linguistic analogues, but not for the phonemic vowel contrast. This confirms reports that the link between musicality and non-native speech sound processing is strongest for tonal contrasts (Delogu et al., 2006, 2010). Adding temporal, pitch and spectral acuity to the model qualified this picture: We had hypothesized that discrimination of tonal contrasts would be predicted by temporal as well as pitch and spectral acuity, whereas discrimination of vowel contrasts would be predicted mainly by spectral acuity. Indeed, for tonal contrasts and pitch contours this prediction bore out, although some effects fell short of significance. This confirms the previously observed role of temporal information, in addition to pitch and spectral information, in the processing of pitch contours, whether embedded in linguistic material or presented on their own (Kempe et al., 2012)—an important finding as researchers often conceive of lexical tones as predominantly involving pitch processing, even though lexical tones often entail changes of pitch and spectral information over time.

Counter to our prediction, we found that, along with spectral acuity, temporal acuity was also a significant predictor for vowel processing. This is surprising because we had carefully controlled for vowel length and metrical stress to exclude temporal information as an additional cue. Still, it is possible that participants were sensitive to subtle durational differences in other segments when trying to discriminate between the Norwegian words. Interestingly, performance with the vowel contrast was also significantly predicted by self-rating in an L3, which suggests that some of the languages participants studied later in life might have provided prior exposure to the /i/ - /y/ vowel contrast, and this experience may have transferred to our stimuli. The effects of prior language experience notwithstanding, our findings suggest that both temporal and spectral acuity were important predictors for the particular phonological contrast used in this study; the lack of a significant correlation of either AMMA scores with temporal acuity might explain why musical aptitude did not predict performance for this particular contrast.

More generally, our findings suggest that temporal, spectral and, to some extent, pitch acuity, underlie processing of a variety of non-native speech sounds. This adds an important facet to our understanding of speech-sound processing in light of approaches that focus on rapid temporal auditory processing as the main sensory component underlying language (Goswami et al., 2002; Tallal, 1980). Moreover, subtle differences in the role of spectral, pitch, and temporal acuity in the processing of different phonemes might account for why links between musicality and non-native phoneme processing sometimes remain elusive (Delogu et al., 2006, 2010). For the tonal contrasts, on the other hand, the residual effects of musical aptitude may reflect shared working memory components (Williamson, Baddeley & Hitch, 2010), in addition to sensory components, due to the somewhat longer duration of these stimuli. In addition, the finding that sensory acuity played a similar role in the processing of tonal contrasts and their non-linguistic analogues challenges the view that linguistic stimuli enjoy a special status with respect to basic sensory processing (Gandour et al., 2000)—a conclusion that needs to be verified through neuro-imaging studies.

The finding that the AMMA rhythm score was a better predictor of non-native tonal and pitch contour processing than the tonal score confirms similar findings with respect to non-native speech sound processing (Nardo & Reiterer, 2009), as well as auditory working memory and reading performance in children (Strait et al., 2011). It appears that information related to repetitiveness, predictability and the sequential nature of sound sequences as measured by the rhythm score has greater diagnostic validity for detecting subtle changes in linguistic stimuli than information related to pitch differences.

A number of studies have conceptualized musicality as musical expertise and suggested that exposure to, and regular practice of music may hone sensory and cognitive abilities, which can subsequently benefit language processing (Marie et al., 2011; Wong et al., 2007). If this were correct one would expect experience with music, measured as number of years of musical training, to exert an independent effect on non-native speech sound processing. Musical training in our sample ranged from 0 to 20 years. Still, no independent effects of length of musical training on non-native speech-sound processing were found. Thus, our findings do not support the notion that musical expertise has independent effects over and above musical aptitude; rather, they suggest that auditory sensory acuity benefits performance both in the musical and the linguistic domain.

Finally, it is worth mentioning that our findings partially replicated the sex difference in non-native speech sound processing observed in previous studies (Bowles et al., 2011; Kempe et al., 2012). This sex difference was assumed to be due to a general male advantage in rapid temporal auditory processing (Wittman & Szelag, 2003) and temporal discrimination tasks (Rammsayer & Troche, 2010), and, thus, should be found for processing of speech sounds that require rapid temporal auditory processing. Indeed, the difference in sensitivity for tonal contrasts and extracted pitch contours pointed towards a male advantage, although the effect fell short of statistical significance in the multiple regression analyses. In further support, a separate analysis for just the tonal and pitch contour stimuli yielded a main effect of sex, $F(1, 101) = 4.7$, $p < .05$. Surprisingly, however, the predicted male advantage was not observed for temporal acuity, $p = .9$; instead, we found a male advantage in pitch acuity, $t(101) = 2.06$, $p < .05$. This suggests that while a male advantage in processing of non-native tonal contrasts seems to be a robust phenomenon, future research may have to explore alternative explanations with respect to the underlying mechanisms.

In sum, our findings further illuminate the link between musical aptitude and non-native speech-sound processing by demonstrating that this link is largely explained by sensory acuity. Our results also suggest that temporal, pitch, and spectral acuity all contribute to the processing of a range of non-native speech sounds.

References

Boersma, P. & Weenink, D. (2011). *Praat: doing phonetics by computer* [Computer program] *Version 5316*, Available: http://wwwfonhumuvanl/praat/. Accessed 2012, May 30.

Bowles, A. R., Silbert, N. H., Jackson, S. R. & Doughy, C. J. (2011). Individual differences in working memory predict second language learning success. Poster presented at the *52nd Annual Meeting of The Psychonomic Society*, Seattle, WA.

Cattell, R. B. & Cattell, H. E. P. (1973). *Measuring Intelligence with the Culture-Fair Tests.* Champaign, IL: Institute for Personality and Ability Testing.

Cutting, J. E. & Rosner, B. S. (1974). Categories and boundaries in speech and music. *Perception and Psychophysics, 16,* 564–570.

Delogu, F., Lampis, G., Belardinelli, M. O. (2006). Music-to-language transfer effect: may melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processes, 7,* 203–207.

Delogu, F., Lampis, G., Belardinelli, M. O. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology, 22,* 46-61.

Duncan, J., Emslie, H., Williams, P., Johnson, R. & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30,* 257-303.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., et al. (2000) A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience, 12,* 207–222.

Gordon, E. E. (1989). *Advanced Measures of Music Audiation.* Chicago: GIA.

Goswami, U., Thompson, J., Richardson, U., Stainthorp, R., Highes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: a new hypothesis. *Proceedings of the National Academy of Sciences USA, 99,* 10911–10916.

Hayes, A. F. & Preacher, K. J. (2013). Statistical mediation analysis with a multicategorical independent variable [White paper]. Retrieved from http://www.afhayes.com

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O. & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience, 23*, 2701–2715.

Nardo, D., Reiterer, S. M. (2009). Musicality and phonetic language aptitude. In: Dogil, G., Reiterer, S. M. (Eds.). *Language talent and brain activity. Trends in applied linguistics* Berlin: Mouton De Gruyter.

Patel, A. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6,* 674–681.

Rammsayer, T. H. & Brandler, S. (2007). Performance on temporal information processing as an index of general intelligence. *Intelligence, 35,* 123–139.

Rammsayer, T., & Troche, S. (2010). Sex differences in the processing of temporal information in the sub-second range. *Personality and Individual Differences, 49,* 923-927.

Slevc, L. R., Miyake, A. (2006). Individual differences in second-language proficiency: *Does musical ability matter? Psychological Science, 17*, 675–681.

Strait, D. L., Hornickel, J. & Kraus, N. (2011) Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behavioral and Brain Functions, 7*.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, *9*, 182-198.

Williamson, V. J., Baddeley, A. D. & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition, 38,* 163-175.

Wing, H. D. (1968). *Tests of musical ability and appreciation: An investigation into the measurement, distribution, and development of musical capacity* (2nd ed.). London: Cambridge University Press.

Wittmann, M., & Szelag, E. (2003). Sex differences in perception of temporal order. *Perceptual and Motor Skills, 96,* 105–112.

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., Belin, P. & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences, 6,* 37-46.