Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Domain-generality of pitch processing: the perception of melodic contours and pitch accent timing in speech

Tuuli H. Morrill,[*#1] J. Devin McAuley,[*2] Laura C. Dilley[#*3], David Z. Hambrick[*4]

[*]*Dept. of Psychology, Michigan State University, USA*

[#]*Dept. of Communicative Sciences and Disorder, Michigan State University, USA*

[1]tmorrill@msu.edu, [2]dmcauley@msu.edu,[3]ldilley@msu.edu,[4]hambric3@msu.edu

ABSTRACT

Background

Whether processing mechanisms are shared between music and speech has been a topic of substantial interest in the past decade. In particular, it is not clear to what extent individuals with pitch processing deficits in music also show these deficits in speech. Prosodic cues, or pitch and timing information in speech, vary across languages, but frequently consist of pitch patterns to convey word-level or sentence level changes in meaning. Studies examining the domain-generality of pitch processing deficits in amusia have yielded mixed results; these studies have primarily investigated the magnitude, direction and rate of pitch changes within a spoken syllable (Patel et al, 2008; Liu, et al, 2010). However, to use pitch information in linguistic processing, listeners must also accurately perceive the timing of pitch changes with respect to syllables. For example, a pitch peak on the second syllable of *digést* (verb) distinguishes the word from *dígest* (noun), in which the pitch peak occurs on the first syllable. In other cases, more subtle changes in peak timing within a syllable can result in distinctive pitch contours; the accurate perception of these prosodic cues is essential to effective speech processing. Since pitch contour processing is also a major component of music perception, perceiving changes in pitch peak timing in speech may rely on domain-general processing.

Aims

The Montreal Battery of Evaluation of Amusia (MBEA) is the standard method for diagnosing congenital amusia. The MBEA is also correlated with measures of musical ability (Peretz et al, 2003).The present study assesses the relationship between performance on the MBEA and the ability to detect changes in the timing of a pitch peak in speech, while controlling for individual differences in cognitive ability. The timing of pitch peaks relative to syllables is used to convey meaning in English and other languages. Thus, the ability to detect fine-grained changes in pitch patterns may be common to both music and speech perception. Here, we manipulated pitch peak timing across two syllables within a nonsense word while maintaining the magnitude, direction and rate of the pitch excursion. The relationship between performance on the MBEA, the speech prosody test, and the battery of cognitive ability measures is assessed with structural equation modeling.

Method

English-speaking participants (n=179) completed a Cognitive Ability Battery that included assessments of working memory capacity (WMC), fluid intelligence (Gf) and crystallized intelligence (Gc), followed by the Montreal Battery of Evaluation of Amusia (MBEA), and a speech prosody test. Tests of WMC included operation span and symmetry span; tests of Gf included matrix reasoning and letter sets; tests of Gc included vocabulary and general information. Subtests of the MBEA considered in the analyses include the Scale, Contour, Interval, Meter, and Rhythm tests. In the speech prosody test, participants were first familiarized with two comparison versions of a nonsense word embedded in a carrier phrase in which a pitch peak occurred with relatively early timing (i.e., on the first syllable of the nonsense word) or relatively late timing (i.e., on the second syllable). Participants then heard four test versions of the phrase, in which the pitch peak occurred with temporal positions intermediate to those of the comparison versions; participants indicated whether this version provided a closer match to the first comparison version or the second comparison version.

Results

Two major steps were involved in the structural equation modeling.

First, we established a measurement model. The final model for the predictor constructs included latent variables representing fluid intelligence and working memory capacity (Gf/WMC), crystallized intelligence (Gc), and music perception (MBEA). Model fit was excellent: $\chi^2(41) = 67.80$, CFI = .96, NFI = .90, RMSEA = .06. We also tested a model with separate Gf and WMC factors, and another model with separate pitch processing and temporal processing factors. However, in both cases, the factors correlated very highly, and improvement in model fit was non-significant; therefore, we modeled a single music perception (MBEA) factor and a single Gf/WMC factor.

Second, we tested for effects of Gf, Gc, and MBEA on a latent variable representing prosody test performance (Prosody). Gf and Gc correlated strongly with each other (r = .63), and each factor correlated moderately and positively with MBEA (rs = .40 and .42, respectively). Nevertheless, only MBEA was a significant predictor of Prosody (β = .55); effects of Gf and Gc on Prosody were near zero (β = .03 and .10, respectively). MBEA accounted for 35.7% of the variance in Prosody; Gf and Gc added less than 1%.

Conclusions

These results indicate that music perception abilities are highly predictive of speech prosody perception abilities. In addition, these effects do not appear to be mediated by intelligence or working memory capacity, as these factors were not predictive of prosody test performance. The prosodic cue that was manipulated in the current study was pitch peaking timing; therefore, the results suggest 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. In music, the perception of a pitch excursion includes assessments of both magnitude (interval) and shape (contour). Interestingly, the Contour and Interval subtests of the MBEA were the greatest predictors of performance on the MBEA and prosody test. If music and speech prosody perception abilities are correlated in this way, people with deficits in music processing may have difficulty with specifically those aspects of speech prosody which require the assessment of intonation pattern alignment with speech structure and content (i.e., syllables, words, and phrases). Since intonation patterns in speech are used to convey meaning, perceptual difficulties with the assessment of pitch peak alignment in amusic individuals may result in deficits in speech processing. The results of this study suggest that the ability to perceive changes in pitch peak timing may be controlled by a domain-general processing mechanism.

Keywords

Intonation, pitch, prosody, timing

REFERENCES

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification, and imitation. *Brain*, 133, 1682-1693.

Patel, A. D., Wong, M., Foxton, J., Lochy, A. & I. Peretz. (2008). *Music Perception* 2 (4), 357–368.

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of Musical Disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, 999, 58-75.