Modular Processing? Phonetic Information Facilitates Speech and Song Imitation

James T. Mantell & Peter Q. Pfordresher

University at Buffalo, The State University of New York, USA

BACKGROUND AND AIMS

The longstanding debate on the nature of music and language is paralleled by discussion about the structure of these systems. The modularity approach, as stated by Peretz and Coltheart (2003), contends that music processing is handled exclusively and without interference from language modules. Peretz and Coltheart proposed a model that illustrates musical perception and production as an outcome of information flow between several processing subcomponents involved in both pitch and temporal organization. According to this model, and by an adaptation by Welch (2005), song and speech processing occur along distinct pathways, but the musical and phonological lexicons can interact. Some music processing, such as tonal encoding, should not be influenced by speech information. Although speech and song are traditionally described as distinct vocal activities, several vocal forms blur the supposed dividing line (e.g. chants and laments, Feld & Fox, 1994). The existence of 'ambiguous' vocal forms suggests that speech and song systems may not be entirely modular and independent. Instead, some aspects of the perception and production of speech and song may be inseparable and highly interactive. In the current work, an intentional imitation paradigm was utilized to investigate the accuracy with which normal individuals imitate the pitch-time trajectories of speech and song target sequences. Based on the modular account of music processing, we developed two predictions. First, imitation of pitch should be better for melodies than sentences because song input has exclusive access to higher level pitch encoding modules. Second, phonetic information should facilitate speech imitation, but not song imitation, because phonology is a necessary component of speech and not necessarily integrated with melody.

METHOD

In all experiments, participants imitated sentence and melody targets that were presented in their original, worded form and as phonetically neutral, wordless sequences where the syllables were synthesized to a 'hum.' Melodies were created based on the pitch-time contour of sentences but were diatonic and isochronous. Experiments 1 and 2 investigated whether or not phonetic information assists or disrupts imitative performance; in experiment 1, the original targets were imitated as heard but in experiment 2, participants imitated all sequences using the neutral syllable "ah." Experiments 3, 4, and 5 assessed the influence of phonetic information when the temporal structure of target sequence was manipulated. In these experiments, participants imitated the temporally altered sequences as they heard them. In experiment 3, speech and song duration were equated. In experiment 4, targets' syllable timing was always speech-like; in experiment 5, syllable timing was isochronous, like melodies. Data analysis consisted of converting the pitch-time information in the vocal recordings to numerical matrices and applying quantitative global accuracy measures. Pitch accuracy was assessed as average absolute error and as correlation between the matched target and imitation vectors; these measures corresponded to absolute pitch matching and relative pitch matching accuracy. Pitch accuracy was calculated by comparing matched pairs of target and imitation sequences after adjusting for differences in production rate. These timing adjustments were used to measure the imitation of overall sequence

ISBN: 1 876346 62 0 © 2010 ICMPC 11

rate. We conducted 2 (sequence type: sentence and melody) x 2 (phonetic content: worded or wordless) repeated measures ANOVAs to assess main effects and interactions.

RESULTS

Prediction one: Partial support for a melodic advantage comes from only one of three measures of imitative accuracy. Analysis of average absolute error revealed a significant main effect of sequence for all five experiments: absolute pitch information was imitated more accurately in melodies than in sentences. Conversely, imitation of relative pitch was not better in melodic imitation except in experiment three when target sequences were equated for overall duration. Importantly, this advantage was for worded melodies only; wordless melodies were no more accurate than worded or wordless sentences. Analysis of timing accuracy revealed that melodies were only better than sentences in experiment two, when all sequences were imitated on the neutral syllable "ah." Again, the advantage was for worded melodies only; imitation of timing in wordless melodies was not more accurate than worded or wordless sentences.

Prediction two: Contrary to expectations, both speech and song sequences benefitted from the inclusion of phonetic information. Furthermore, the phonetic benefit may even be stronger for song than speech imitation. Both absolute and relative pitch were imitated more accurately in worded than wordless melodies in experiments 1, 3, 4, and 5; the only experiment that did not produce a phonetic advantage for pitch imitation was experiment 2. Analysis of timing accuracy indicated a phonetic advantage in experiment 1. In experiment 2, imitation of melodic timing did not change, but imitation of worded sentence timing worsened, suggesting that speech imitation may be somewhat more sensitive to mismatch in the perception and production of phonetic information.

CONCLUSIONS

The results strongly suggest that speech does not gain exclusive benefit from phonetic information; song imitation was facilitated when phonetic information was presented in the target and utilized in the production. These data do not favor a strong modular account of speech and song systems; the superiority of imitations of worded melodies over other target types suggests strong interaction between phonology, phonation and some aspects of pitch and time processing. Future experiments using intentional vocal imitation could further illuminate the relationship between music and language systems by focusing on the role of sequence memory in imitative performance.

REFERENCES

Feld, S., & Fox, A. A. (1994). Music and language. *Annual Review of Anthropology, 23*, 25-53.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 688-691.

Welch, G. F. (2005). Singing as communication. In D. Miell, R. MacDonald, & D. J. Hargreaves (Eds.), *Musical communication* (pp. 239-260). Oxford: Oxford University Press.

TOPIC AREAS

Music and Language