ELSEVIER

Cognition 81 (2001) B11–B20

COGNITION

www.elsevier.com/locate/cognit

Brief article

The effect of harmonic context on phoneme monitoring in vocal music

E. Bigand[a,*], B. Tillmann[b], B. Poulin[a], D.A. D'Adamo[a], F. Madurell[c]

[a]*Université de Bourgogne, Dijon, France*
[b]*Dartmouth College, Hanover, NH, USA*
[c]*Université Paris IV-Sorbonne, Paris, France*

Received 27 September 2000; received in revised form 6 December 2000; accepted 31 January 2001

Abstract

The processing of a target chord depends on the previous musical context in which it has appeared. This *harmonic priming* effect occurs for fine syntactic-like changes in context and is observed irrespective of the extent of participants' musical expertise (Bigand & Pineau, *Perception and Psychophysics*, *59* (1997) 1098). The present study investigates how the harmonic context influences the processing of phonemes in vocal music. Eight-chord sequences were presented to participants. The four notes of each chord were played with synthetic phonemes and participants were required to quickly decide whether the last chord (the target) was sung on a syllable containing the phoneme /i/ or /u/. The musical relationship of the target chord to the previous context was manipulated so that the target chord acted as a referential *tonic chord* or as a congruent but less structurally important *subdominant chord*. Phoneme monitoring was faster for the tonic chord than for the subdominant chord. This finding has several implications for music cognition and speech perception. It also suggests that musical and phonemic processing interact at some stage of processing. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Priming; Expertise; Phoneme monitoring; Vocal music; Modularity

* Corresponding author. LEAD CNRS, Faculte des Sciences, 6 Boulevard Gabriel, 21000 Dijon, France. Fax: +33-380-395-767.

E-mail address: bigand@satie.u-bourgogne.fr (E. Bigand).

PII: S0010-0277(01)00117-2

1. Introduction

A previous context primes the processing of related events. Priming effects occur for a large set of stimuli and have been largely documented in the language domain. In the case of spoken language, phoneme monitoring studies have shown faster processing for phonemes occurring in words rather than in non-words (Cutler, Melher, Norris, & Segui, 1987), in high frequency target words (Dupoux & Melher, 1990), in highly predictable words (Dell & Newman, 1980; Eimas & Nygaard, 1992; Morton & Long, 1976), in words that match the semantic focus of the sentence (Cutler & Fodor, 1979), or in words that follow semantic associates (Frauenfelder & Segui, 1989). In the music domain, harmonic priming studies demonstrated that the processing of a target chord (say a G major chord) is faster and more accurate when a prime belongs to the same musical key (a C major chord) than when it does not (an F# major chord) (Bharucha & Stoeckig, 1986, 1987; Tekman & Bharucha, 1998). Harmonic priming also occurs in a larger musical context. The processing of target chords ending eight-chord sequences was facilitated when the target acted as a referential tonic chord in the key context rather than as a congruent, but less referential subdominant chord (Bigand, Madurell, Tillmann, & Pineau, 1999; Bigand & Pineau, 1997). This finding suggests that harmonic priming does not only occur from chord to chord but derives from various levels of the musical structure.

Up to now, linguistic and musical priming have been investigated separately. There are, however, several reasons to assess whether the one structure may prime the processing of the other (Besson, 1998). First, most of everyday life involves vocal music, and it is thus very common to process both structures simultaneously. Second, because harmonic and linguistic structures are entwined in a non-arbitrary way in vocal music, we may expect both structures to be integrated during perceptual processing or in memory. The dependence or independence of music and language processing in vocal music remains nevertheless a matter of debate. Some research provided strong evidence in favor of independence of processing (Besson, Faïta, Peretz, Bonnel, & Requin, 1998), while other research suggested that both structures are integrated (Serafine, Crowder, & Repp, 1984). Finally, the potential influence of one structure on the processing of the other has implications for models specific to each domain. In the present study, we were notably interested in assessing whether the processing of harmonic structures occurs in an automatic way so that it might bias the processing of a non-musical feature to which participants were asked to pay explicit attention. In addition, we also suggest that the use of a linguistic task in music cognition experiments has methodological advantages for evaluating the influence of musical expertise.

The present study focuses on the effect of musical structure on the processing of the most basic linguistic unit: the phoneme. Eight-chord sequences sung with sampled phonemes were presented to participants. Their task was to indicate as quickly as possible whether or not the consonant–vowel syllable sung on the last chord (target chord) contained the phoneme /i/ or /u/. The critical point consisted of manipulating the harmonic relationship of the target chord to the previous context. The rationale of the manipulation comes from previous harmonic priming studies.

Western music contains keys (e.g. C major, D major, F minor) that are made of seven chords, each constructed on a different scale degree of the key. A critical feature of the Western musical system is that chords have different syntactic-like functions inside a key. Chords built on the first, fifth, and fourth scale degrees (referred to as the *tonic*, *dominant* and *subdominant* chords, respectively) have a more central syntactic function than other chords of the key, the tonic chord being the most referential event of the key. Harmonic priming is supposed to be strongly related to the musical function of the chord in the previous key context. Chords that do not belong to the key context (Bharucha & Stoeckig, 1987; Tillmann, Bigand, & Pineau, 1998), or that are less referential in the key context (such as a subdominant chord compared to a tonic chord) are less primed by the context resulting in slower processing (Bigand et al., 1999; Bigand & Pineau, 1997).

In the present study, the harmonic function of the target chord was manipulated in such a way that the target acted as a tonic chord in the *related condition* and as a subdominant chord in the *congruent but less related condition* (Fig. 1). In both related and less related conditions, the local harmonic relationship between the target chord and the immediately preceding chord was held constant and the target

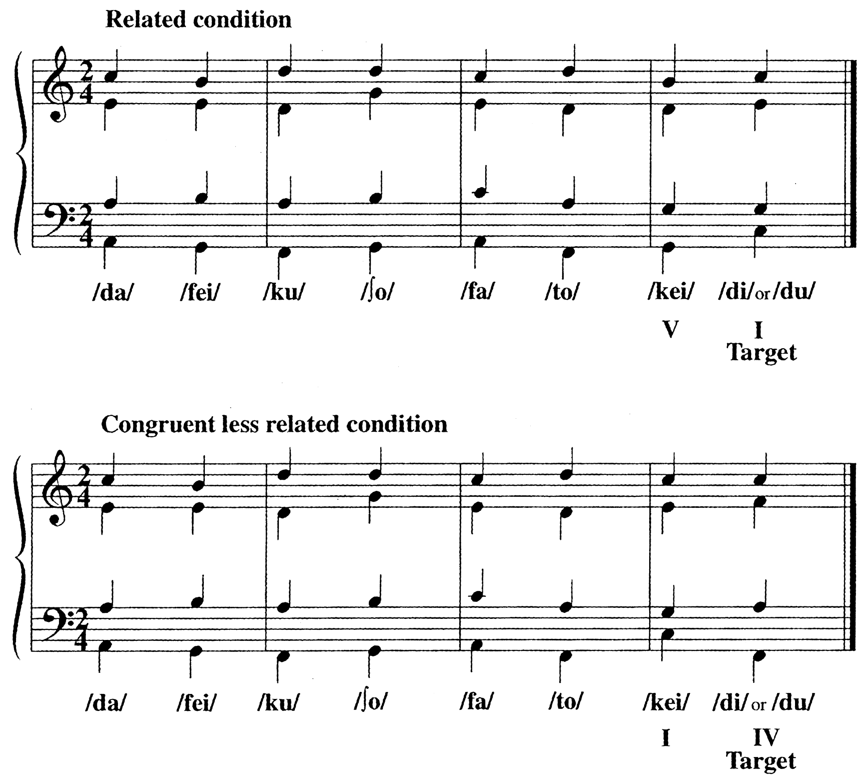

Fig. 1. One example of the sung chord sequences used in the experiment.

chord never occurred in the previous context. Phoneme monitoring was expected to be faster in the *related condition* than in the *congruent but less related condition*. In addition, a phoneme monitoring task permitted further investigation of the contribution of musical expertise to harmonic priming. Up to now, harmonic priming has been demonstrated by requiring participants to perform a perceptual task on the target chord that was more familiar to musicians than non-musicians. Even if harmonic priming has been reported for both groups of participants, overall musicians were faster and more accurate than non-musicians. The use of the phoneme monitoring task permits better assessment of the influence of musical expertise because both groups of participants are likely to have the same expertise with phoneme processing.

2. Method

2.1. Participants

Forty students participated in the experiment: 19 students with no formal musical training or any practice with a musical instrument (referred to as “non-musicians”), and 21 graduate students of the national music conservatories of Troyes, Dijon, and Grenoble (referred to as “musicians”).

2.2. Materials[1]

Twelve musical sequences (one per major musical key) containing six chords were defined for the purpose of the experiment. For each of them, two types of ending (seventh and eighth chords) were added, resulting in 24 sequences: in one (*related condition*), the eighth chord was a tonic chord, whereas in the other (*congruent but less related condition*) it was a subdominant chord (see Fig. 1 for one example). The global context, created by the first six chords, defined the function of the corresponding target chord. In both contexts, the local harmonic relationship between the seventh chord and the eighth chord (the target) defined an authentic cadence[2] (e.g. G-C or C-F). These chord sequences were composed in such a way that the target chord never occurred in the previous context of the sequence. This construction allowed us to control a potential confound between the change in the target’s musical function and its frequency of occurrence in the context.

[1] Sound examples of the stimuli may be found on the WEB at the following address: http://www.u-bourgogne.fr/LEAD/people/bigand/choralVocal.htm

[2] An authentic cadence is a succession of two major chords that creates a clear impression of a definitive ending. The roots of the second chord are five semitones above or seven semitones below the root of the first one (G-C).

2.3. Apparatus

The eight-chord sequences were played with sample voice sounds by the VocalWriter singing music software, version 1.0 (Cecys, 1998). This software permits the assignment of a specific syllable to each chord and the playing of the four notes of a chord with the same syllable. Twenty-four consonant–vowel syllables that were very different from French syllables were used[3]. The syllables /di/ and /du/ were always assigned to the target chord. The remaining 22 syllables were randomly assigned to the first seven chords of the sequences. The phonemes /i/ and /u/ occurring in the syllables /di/ and /du/ were chosen as the target phonemes because they were the most easily distinguishable phonemes available in VocalWriter. The sound stimuli were captured by SoundEdit Pro software with CD quality and the experiment was run on PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). The tempo of the sequences was 96 quarter tones per minute (i.e. 625 ms per chord). The duration of the target phoneme was limited to 2 s and the sound disappeared as soon as the participants had responded.

2.4. Procedure

Participants were first trained with isolated sung chords to quickly discriminate between the syllables /di/ and /du/. In the experimental phase, the eight-chord sequences were presented and participants were asked to quickly decide whether the syllable sung on the last chord contained the phoneme /i/ or /u/. They were alerted by a feed-back signal if they gave an incorrect response. Crossing the Harmonic Relationship (related versus less related) and the Target Type (phoneme /i/ versus /u/) produced four possible versions for each of the 12 musical sequences. Each participant thus heard 48 chord sequences presented in a random order.

3. Results

The high percentages of correct responses (98 and 96% for musicians and non-musicians, respectively) demonstrated that French participants had no difficulty discriminating between the English phonemes /i/ or /u/, irrespective of musical expertise. Despite this high level of performance, a 2 (Target Type) $\times$ 2 (Harmonic Relationship) $\times$ 2 (Musical Expertise) ANOVA was performed on accuracy data. There was only one significant effect: the percentage of correct responses was higher for the tonic target chord (98.2%) than for the subdominant target chord (96.8%) ($F(1, 38) = 7.28$, MSE $= 0.1533$, $P < 0.01$).

Averaged correct response times are presented in Table 1. There was a main effect

---

[3] These consonant–vowel phonemes are represented in VocalWriter software with the following notation: dIY; dUW; dAA; dEH; dOH; fAA; fEH; fIY; fOH; fUW; kAA; kEH; kIY; kOH; kUW; SHAA; SHEH; SHIY; SHOH; SHUW; tEH; tIY; tOH; tUW, which correspond to the following phonetic transcriptions: /di/; /du/; /da:/; /de:/; /d*o*:/; /fa:/; /fe/; /fi/; /f*o*/; /fu/; /ka:/; /ke/; /ki/; /k*o*/; /ku/; /sa:/; /se:/; /si/; /s*o*/; /su/; /te:/; /ti/; /t*o*/; /tu/.

Table 1
Correct response times as a function of harmonic relationship, target type and musical expertise (standard errors in parentheses)

Phoneme	Correct response times			
	Musicians		Non-musicians	
	Tonic	Subdominant	Tonic	Subdominant
/i/	501.97 (40.90)	532.18 (44.89)	562.37 (44.12)	606.77 (56.60)
/u/	531.10 (49.33)	580.64 (52.52)	624.80 (56.44)	644.50 (50.75)

of Target Type, with shorter response times for the phoneme /i/ ($F(1,38) = 12.40$, $MSE = 6352$, $P < 0.001$), and a main effect of Harmonic Relationship ($F(1,38) = 14.31$, $MSE = 3606$, $P < 0.001$): phoneme monitoring was faster when the phoneme was sung on the tonic chord than when it was sung on the subdominant chord. There were no other significant effects. Musicians tended to respond faster than non-musicians (536 versus 609 ms), but this difference was not significant ($F(1.38) = 1.17$). Interestingly, at the end of the experiment, none of the musicians reported noticing that the sequences ended either on a tonic or a subdominant chord. An additional ANOVA performed with chord sequences as a random variable confirmed the main effects of Harmonic Relationship and Target Type: phoneme monitoring was faster in the related condition ($F_2(1,22) = 12.76$, $P < 0.01$), and faster for the phoneme /i/ ($F_2(1,22) = 19.77$, $P < 0.001$).

4. Discussion

The present experiment demonstrates an effect of harmonic context on the processing of sung phonemes: phoneme monitoring was faster for phonemes sung on chords harmonically related to the context. This finding has several implications for research on music cognition, speech perception, and for the issue of modularity of language and music processing.

First, the observed priming effect furthers our understanding of music cognition as it provides some evidence that harmonic priming involves a sophisticated cognitive component that occurs in an automatic way. Participants were not asked to pay explicit attention to the musical structure of the sequences (the task only required the participants to focus on the last syllable), but their performance demonstrated that the harmonic function of the sung target chord was nevertheless processed. It is worth noting that in the present experiment, the target never occurred in the context. One consequence is that the processing of the musical function of the target is rendered more difficult. Indeed, in order to differentiate subtle changes in the function of the target (tonic versus subdominant), the key of the context must be inferred first. This inference is more difficult when neither the tonic chord nor the subdominant chord occurs in the context. The fact that the facilitation effect continues to be observed for the tonic rather than the subdominant chord emphasizes the strength of the cognitive process involved in harmonic priming.[4] A further implication of the present findings concerns the influence of musical expertise on music cognition. The phoneme monitoring task allows the comparison of musicians and non-musicians using an experimental task that is equally familiar to both groups of participants. The fact that with this task non-musicians behave exactly as musical experts do is impressive and sheds new light on the weak influence of musical expertise on the processing of harmonic structures. Taken in combination, these findings point to the robustness of the cognitive processes into which harmonic priming taps (see Tillmann, Bharucha, & Bigand, 2000). They also suggest that the musical abilities of the large audience are likely to be generally underestimated in common social belief.

The second main implication of the present study concerns research on speech perception. Phoneme monitoring tasks have been used to investigate several issues of speech processing (see Connine & Titone, 1996 for a review). To the best of our knowledge, phoneme monitoring has never been studied in a musical context. Although some factors influencing phoneme monitoring seem to be clearly specific to linguistic processing (i.e. factors related to lexical, semantic and syntactic effects) others (such as prosody) may tap into more general cognitive processes. The similarity between prosodic aspects of language and some features of music has often been noticed. In a recent study, Patel, Peretz, Tramo, and Labreque (1998) provided evidence that both structures are likely to share neural resources. The effect of harmonic context on phoneme monitoring in our study may be understood along these lines. Prosody studies have demonstrated that faster phoneme monitoring occurs for stressed rather than unstressed target words (Cutler, 1976; Cutler & Foss, 1977; Shield, McHugh, & Martin, 1974), for targets in sentence accent positions (Mehta & Cutler, 1988) and for targets predicted to be in stressed positions (Pitt & Samuel, 1990). This finding is generally interpreted as evidence that stress helps to focus attentional resources on specific moments of the sentence, allowing faster processing of target phonemes (Cutler, 1976; Weismer & Ingrisanno, 1979). In a similar way, music displays several accents whose common function is to capture listeners' attention during the unfolding of a musical piece (Bigand, 1997; Jones, 1987; Jones & Boltz, 1989). Stable events occurring in Western harmonic hierarchies define harmonic accents. Harmonic accents may influence phoneme monitoring in vocal music as prosodic cues influence phoneme monitoring in speech

[4] It might be argued that the transition of harmonic functions between the sixth and seventh chords is more unusual (but legal) in the less related (II → I) than in the related (II → V) condition. The transition II → I may potentially create some local surprise resulting in slowing down of the processing of the subdominant targets in the less related condition. However, it should be noticed that similar priming effects are observed when the transition between the sixth and seventh chords in the less related condition is a very usual V → I progression (Bigand & Pineau, 1997). Moreover, the same size of priming is observed when in the same set of chord sequences the II → I transition is replaced by a IV → I transition (Poulin, Bigand, D'Adamo, Madurell, & Tillmann, 2000). Finally, harmonic priming continues to be observed in 14-chord sequences even when the last seven chords of the related and less related sequences are held constant (see Bigand et al., 1999). These findings suggest that harmonic priming is not linked to the local presence of more or less usual transitions.

perception. In the present experiment, tonic target chords might act as stronger stressed events than subdominant chords, resulting in faster processing for the phonemes sung on the tonic rather than the subdominant chord.

Finally, the present study has some implications for the debate on the modularity of music and language processing. Independent effects were reported by some authors (Besson et al., 1998 for the processing of semantics and music; Palmer & Kelly, 1992 for the processing of prosody and meter), while other research suggests the integration of structures (Serafine et al., 1984 for text and melody) or the overlap of representation at certain neural levels (Patel et al., 1998 for prosody and musical patterns). Recent studies directly dealing with the processing of phonetic and musical sounds have demonstrated that the processing of phonetic and musical information involves spatially distinct neuronal populations. The processing of phonetic information was associated with higher activity in the left auditory cortex and a reverse phenomenon was found for musical chords (Tervnanimi et al., 1999, 2000).

Our present study showed that manipulating the harmonic structure of a musical sequence influences the processing of phonemes even when participants are not asked to pay attention to the music. This suggests that the processing of phonetic and musical sounds is not independent, but interacts at some level of processing. It may be argued, however, that both processes are independent with the processing of the harmonic function occurring before the processing of the phonetic information. A delay in the former process (caused by the occurrence of a less related subdominant chord) would delay the latter one. This interpretation is questionable for two reasons. First, the processing of semantics was shown to occur before the processing of music in lyrics (Besson et al., 1998). Second, we may wonder why the processing of an attentionally non-focused feature (music) should occur before the processing of a focused one (phoneme).

Acknowledgements

This research was supported by the International Foundation for Music Research.

References

Besson, M. (1998). Meaning, structure, and time in language and music. *Current Psychology of Cognition*, *17*, 921–950.

Besson, M., Faïta, F., Peretz, I., Bonnel, A.-M., & Requin, J. (1998). Singing in the brain: independence of lyrics and tunes. *Psychological Science*, *9*, 494–498.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy. *Journal of Experimental Psychology: Human Perception and Performance*, *12*, 403–410.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: spreading activation or overlapping frequency spectra? *Perception and Psychophysics*, *41*, 519–524.

Bigand, E. (1997). Perceiving musical stability: the effect of tonal structure, rhythm and musical expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 808–812.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal

organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 184–197.
Bigand, E., & Pineau, M. (1997). Context effects on musical expectancy. *Perception and Psychophysics*, *59*, 1098–1107.
Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: an interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers*, *25*, 257–271.
Connine, C., & Titone, D. (1996). Phoneme monitoring. *Language and Cognitive Processes*, *11*, 635–645.
Cutler, A. (1976). Phoneme-monitoring as a function of preceding intonation contour. *Perception and Psychophysics*, *20*, 55–60.
Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition*, *7*, 49–59.
Cutler, A., & Foss, D. J. (1977). On the role of sentence stress in sentence processing. *Language and Speech*, *20*, 1–10.
Cutler, A., Melher, J., Norris, D., & Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, *19*, 141–177.
Dell, G. S., & Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning and Verbal Behavior*, *19*, 608–623.
Dupoux, E., & Melher, J. (1990). Monitoring the lexicon with normal and compressed speech: frequency effects and the prelexical code. *Journal of Memory and Language*, *29*, 316–335.
Eimas, P. D., & Nygaard, L. C. (1992). Contextual coherence and attention in phoneme monitoring. *Journal of Memory and Language*, *31*, 375–395.
Frauenfelder, U. H., & Segui, J. (1989). Phoneme monitoring and lexical processing: evidence for associative context effects. *Memory and Cognition*, *17*, 134–140.
Jones, M. R. (1987). Dynamic pattern structure in music: recent theory and research. *Perception and Psychophysics*, *41*, 621–634.
Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491.
Mehta, G., & Cutler, A. (1988). Detection of target phonemes in spontaneous and read speech. *Language and Speech*, *31*, 15–156.
Morton, J., & Long, J. (1976). Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, *15*, 43–51.
Palmer, C., & Kelly, M. H. (1992). Linguistic prosody and musical meter in song. *Journal of Memory and Language*, *31*, 525–542.
Patel, A., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain & Language*, *61*, 123–144.
Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attention to speech. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 564–573.
Poulin, B., Bigand, E., D'Adamo, D., Madurell, F., & Tillmann, B. (2000, May). *Does musical expertise influence the processing of harmonic structure?* Poster presented at The Biological Foundations of Music, New York Academy of Sciences Conference, New York.
Serafine, M. L., Crowder, R. G., & Repp, B. H. (1984). Integration of melody and text in memory for song. *Cognition*, *16*, 285–303.
Shield, J. L., McHugh, A., & Martin, J. G. (1974). RT to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology*, *102*, 250–255.
Tekman, H. G., & Bharucha, J. J. (1998). Implicit knowledge versus psychoacoustic similarity in priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 252–260.
Tervnanimi, M., Kujala, A., Alho, K., Virtanen, J., Ilmoniemi, R. J., & Näätänen, R. (1999). Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic (MEG) study. *Neuroimage*, *9*, 330–336.
Tervnanimi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijen, T. L., & Näätänen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Human Brain Mapping*, *10*, 74–79.
Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of music: a self-organizing approach. *Psychological Review*, *107*, 885–913.

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effect of local and global contexts on harmonic expectancy. *Music Perception*, *16*, 99–118.

Weismer, G., & Ingrisanno, D. (1979). Phrase-level timing patterns in English: effects of emphatic stress location and speaking rate. *Journal of Speech and Hearing Research*, *22*, 516–533.