Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 37–47

PERSISTENCE OF PERFORMANCE DETAILS IN MUSIC AND SPEECH

Melissa K. Jungers
Ohio State University, Newark, USA

What aspects of music and speech are retained in memory? How do remembered performance details influence future performances? This paper focuses on memory for performance details in music and speech and the influence of these elements from perception to performance. Listeners form a memory for a sentence or melody that includes timing and intensity details. These details then influence performance. Musicians persist in the tempo of a melody they have just heard. They also incorporate details of timing and intensity into their subsequent performances. Speakers persist in the rate of sentences when they produce similar sentences. As in music, this persistence extends beyond the global dimension of rate.

Introduction

The orchestra finishes triumphantly and the final notes reverberate in the concert hall. The audience members applaud and rise to their feet. What is left of that ephemeral sound? What part of the performance will the listeners keep? Will they walk away humming the tune? Research suggests listeners form a memory for the performance that includes not only the melody, but also the more subtle dynamics, timing, and nuances of the music. This ability to remember more than the basic categorical information is not specific to music: a similar ability to remember performance details is found in the domain of speech. In both domains, these acoustic details are part of the representation in memory for the melody or sentence. When producing a new sentence or melody, the listener is influenced by what was just heard. Thus, the music that seems to float away actually becomes part of the listener's memory and influences how the listener performs in the future.

The focus of this paper is on this influence of acoustic details from perception to performance. Pianists persist in the tempo of a melody they have just heard when they perform a similar melody. This persistence is not restricted to tempo. Musicians also retain information about meter and timing and incorporate these details into their subsequent performances. This persistence effect is also found in the domain of speech. Speakers persist in the rate of sentences when they produce similar sentences. As in music, this persistence extends beyond the global dimension of rate.

Correspondence concerning this article should be addressed to Melissa K. Jungers, e-mail: jungers.2@osu.edu

Musical details in performance and perception

Music is composed with specific notes and rhythms, but performers add acoustic features that are not in the notation. For example, performers may lengthen notes or increase accents. Such subtle acoustic variations are known as performance "expression" and they help listeners to differentiate two performances (Palmer, 1997). These musical nuances relate to the musical structure in an ordered way. Western tonal music is organized by meter and grouping principles (Cooper & Meyer, 1960, Lerdahl & Jackendoff, 1983). Meter refers to the alternation of strong and weak beats. For example, a march notated in 2/4 time will have an alternating strong-weak beat pattern while a waltz in 3/4 time will follow a strong-weak-weak beat pattern. Grouping refers to pitch relationships and rhythmic patterns (Lerdahl & Jackendoff, 1983; Cooper & Meyer, 1960). Meter and grouping are both arranged in a hierarchy, with smaller pitch or rhythmic events nested within larger events.

The fine-grained acoustic details in performance are often tied to this musical structure. For example, performers typically mark the ends of musical phrases by a decrease in tempo and dynamics (Henderson, 1936). This performance nuance is known as phrase-final lengthening and it highlights the importance of the phrase in the musical hierarchy (Lerdahl & Jackendoff, 1983; Palmer, 1996; Palmer, 1997). Performers also use performance features to mark meter. They perform with increased duration, more legato articulation, and louder accents on events that align with metrically strong beats than events that align with weak beats (Sloboda, 1983, 1985). Do these acoustic details interact in performance expression? In a study of musical accents associated with meter, rhythmic grouping, and melodic accents, meter and rhythm independently influenced performance expression (Drake & Palmer, 1993). However, the influence of melodic accent on performance expression depended on the context (Drake & Palmer, 1993).

The connection between performance expression and music structure also influences listeners' perceptions. In one study, listeners heard performances that contained one or more performance cues and judged the intended meter (Sloboda, 1985). The listeners relied on articulation cues and loudness to choose the intended meter, although some performers did not use loudness to indicate meter (Sloboda, 1985). Listeners base their judgments on musical expectations. For example, listeners had difficulty detecting a computer-lengthened event that occurred before a long duration in a simple rhythm (Drake, 1993). This perceptual error occurred in the same place in the music at which performers often lengthen events (Drake & Palmer, 1993). Also, when a computer-generated performance contained a lengthened event at a structurally-expected location, listeners were less likely to detect it (Repp, 1992). In another task, listeners judged how a probe beat fit into a metrical context. Their judgments revealed their implicit knowledge of metrical structure (Palmer & Krumhansl, 1990). Thus, both performers and perceivers relate fine-grained performance details to musical structure.

Musical details in memory

Performance details shape performance and perception, but are these details remembered by the listener? When the audience leaves the concert hall humming, what aspects of the performance have they retained in memory? Early perceptual research focused on listeners' abilities to recognize a tune as the same even when performed by a different instrument or at a different

tempo. This idea of perceptual constancy suggests that a basic version of the music is retained, stripped of the timbre, tempo, and other performance details (Dowling & Harwood, 1986; Large, Palmer, & Pollack, 1995). Thus, a normalized version of the melody remains (Large et al., 1995).

Perhaps listeners do not have memories for performance details because these features are used only to form the basic pitch and rhythmic categories, but are not retained in memory (Raffman, 1993). Raffman (1993) points to research by Siegel and Siegel (1977), that shows trained musicians do not accurately detect small pitch differences within categories. If trained musicians have perceptual difficulty with fine-grained pitch differences, surely the average listener does not retain acoustic details in memory.

However, there is evidence that listeners do remember performance details. In a study of memory for performance details, listeners with and without musical training were familiarized with one of two performances of the same short musical excerpt (Palmer, Jungers & Jusczyk, 2001). These performances contained the same pitches and rhythmic patterns, but differed in articulation, intensity, and interonset interval cues. At test, listeners heard both the original performance from familiarization as well as a different performance of the same melody. Listeners were required to identify the performance they heard at familiarization. Even though the pitch and rhythm categories in the two performances were the same, listeners recognized the performance from familiarization (Palmer et al., 2001).

The Palmer et al. (2001) study demonstrated memory for music performance details in adults. Listeners with and without formal music training could remember and differentiate performances based on fine acoustic details, but this result may be due partly to years of exposure to Western tonal music. To address whether musical acculturation is necessary for memory of musical details, Palmer et al. (2001) tested 10-month-old infants for performance memory using the same melodies. Infants were first familiarized with one performance of each melody. They were then tested with a head-turn preference procedure (Kemler Nelson et al., 1995) on the original and different performances of the same melodies. Infants oriented longer to the familiar performances during test, suggesting even infants retain acoustic cues for performances in memory (Palmer et al., 2001).

In addition to articulation details, musicians retain performance tempi for long periods of time. For example, musicians can perform an entire movement of a symphony at the same tempo as previous performances, with very little variability (Clynes & Walker, 1986; Collier & Collier, 1994). This ability to retain musical timing is not limited to those with musical training. In one study, nonmusicians reproduced popular songs from memory at tempi very close to the original tempo (Levitin & Cook, 1996). Also, the participants showed wide tempo variability for songs that did not have a standard original tempo (Levitin & Cook, 1996).

Speech details in production and perception

Music is not the only domain in which subtle performance variations are produced and perceived. The element of speech that includes these performance details is known as prosody. Informally, prosody is the way something is said. Prosody is both a structure that organizes sound and the suprasegmental features of speech including pitch, timing, and loudness (Cutler, Dahan, & von Donselaar, 1997). Prosody is also described as the "stress, rhythm, and intonation in spoken sentences" (Kjelgaard & Speer, 1999).

The prosody of a sentence, including word duration, timing, and intonation, can influence the listeners' interpretation of meaning. Prosody helps to disambiguate grammatically ambiguous sentences. In one experiment, listeners had to guess the meaning of ambiguous sentences read by four speakers (Lehiste, 1973). Listeners relied on timing and intonation cues and were better than chance for 10 of the 15 sentences (Lehiste, 1973). In another study, listeners heard syntactically ambiguous sentences with prosodic emphasis on different words, such as "They are FRYING chickens" and "They are frying CHICKENS" (Speer, Crowder, & Thomas, 1993). When listeners paraphrased the meaning for each sentence, their interpretations revealed the influence of the prosodic emphasis (Speer et al., 1993). The acoustic details associated with a speaker's voice can also aid sentence interpretation (Nygaard & Pisoni, 1998). Listeners were familiarized with isolated words produced by ten speakers. When they were later tested for intelligibility of novel words in noise, they better identified words spoken by a familiar voice than a new voice (Nygaard & Pisoni, 1998).

How does prosody relate to syntax: the grammatical rules for putting words together? In music, there is a clear connection between performance expression and musical structure, but the connection between prosody and syntax is not as straightforward. Although prosody often marks syntax, the relationship between these aspects of speech is not isomorphic (Cutler et al., 1997). Prosody has its own hierarchical structure (Beckman, 1996). Prosody helps to disambiguate syntax (Price et al., 1991). In one study, listeners judged the point at which recorded sentences switched from one ear to the other (Wingfield & Klein, 1971). The sentences contained a phrase that matched or did not match the intonation of the sentence. The listeners relied on both syntactic form and prosodic pronunciation to determine the switching point. Wingfield and Klein (1971) argued that syntax is the primary cue for sentence segmentation, although prosody helps to mark this syntax.

Although past research indicated that listeners use prosody to interpret syntactically ambiguous sentences (Lehiste, 1973, Lehiste et al., 1976), several recent papers questioned the generalizability of this effect outside of the laboratory. In one study, judges rated the intended meaning of syntactically ambiguous sentences produced by trained and untrained speakers (Allbritton, McKoon, & Ratcliff, 1996). The trained speakers had amateur or professional experience in acting or broadcasting. The speakers produced these sentences by reading them within two different passages that clarified the meaning. Untrained speakers and trained speakers who were unaware of the ambiguity did not disambiguate the sentences, according to the judges' ratings. Only the productions by trained speakers who were informed of the ambiguity were judged to disambiguate the meaning (Allbritton et al., 1996).

Another study of natural speech pitted passage context against sentence prosody (Fox Tree & Meijer, 2000). To create the stimuli, speakers read and memorized a short passage and then produced it. The middle sentence was then replaced with a sentence whose prosodic cues fit or did not fit the context of the passage. Listeners heard the recreated passages and chose the intended meaning of each passage. Listeners' choices were based on the context and not the prosody of the embedded sentence. The authors cited this as evidence that prosodic cues are not useful for syntactic disambiguation in a conversational context (Fox Tree & Meijer, 2000). However, there are several concerns with the experiment that make this conclusion less clear. For example, the stimuli were created by speakers whose task was to memorize and produce the passages verbatim. Although this production method is more natural than reading, the speakers may not have been using the full range of prosodic cues since their focus was to memorize and repeat the passage. Also, listeners could rely on the first sentence alone to interpret the passage.

Additionally, listeners were not instructed to use prosody (or even the middle sentence) to make their decisions. Although the debate about the use of prosody in syntactic disambiguation continues, there is evidence that prosody interacts with syntax (Wingfield & Klein, 1971; Lehiste, 1973).

Memory for prosody

Listeners have an amazing ability to understand speech under many conditions. They understand words spoken by children, men, and women, even though the vocal range is quite different for these groups. They understand speakers with unfamiliar accents. This human ability to understand spoken language with widely varying acoustic properties led early researchers to look for normalization processes. The idea behind normalization is that listeners form a representation of speech that lacks prosodic details (Pisoni, 1997). Thus, according to this view, timing and intonation are not part of the memory for a sentence.

More recent studies suggest that prosody is retained in our memory for language. Sentences that are presented with the same prosody at learning and test are recognized more accurately than sentences with different prosody (Speer et al., 1993). Also, listeners use prosodic cues to remember syntactically ambiguous sentences (Speer et al., 1993). Extralinguistic information, such as talker identity and talker rate, help listeners to identify words presented previously (Bradlow, Nygaard, & Pisoni, 1999). In addition, the presentation rate influences listeners' memory abilities for items produced by different speakers. Listeners more accurately recognize items presented at the same rate from familiarization to test than items presented at different rates (Nygaard, Sommers, & Pisoni, 1995). Prosody is incorporated into the memory for language.

Persistence in music

When musicians play in an ensemble and trade the melody from instrument to instrument, how does the performance of one player influence the performances of the others? Musicians include expressive nuances in their performances that are not notated in the musical score. Further, memory for music includes these details. How do these fine-grained performance details influence future performances?

One aspect of music that may persist from performer to performer is tempo. In one study, pianists were instructed to perform one melody at a particular tempo and then a second melody at either a slower or a faster tempo (Cathcart & Dawson, 1928). Pianists then played the original melody and tried to reproduce the original tempo, but their tempi drifted in the direction of the intervening performance tempo (Cathcart & Dawson, 1928). In a review of many studies with tasks as varied as color perception and weight lifting, Warren (1985) found a more general trend to explain this drift. Each domain showed a perceptual homeostasis so that perceivers' criteria shifted according to the current environmental conditions (Warren, 1985).

More recent research found evidence for persistence of musical tempo (Jungers, Palmer & Speer, 2002). Trained adult pianists first sight-read two melodies at their preferred rate. On each of the following trials, the pianists heard a computer-generated melody and then performed a similar melody. The pianists were not instructed to perform at a particular tempo. The compu-

ter-generated melodies (prime melodies) were blocked by fast (300 ms per eighth-note beat) or slow (600 ms per eighth-note beat) tempo. Pianists performed slower following the slow prime melodies than the fast prime melodies. However, their performances did not demonstrate simple imitation of the performance tempi they had just heard. Instead, the pianists' tempi reflected a drift away from their preferred tempo toward the prime melody tempo. Thus, the tempo of the prime melodies influenced the pianists' performances of the target melodies.

Although Jungers et al. (2002) demonstrated persistence of tempo in music, their study did not address other acoustic dimensions. In a set of experiments, Jungers (2003) examined whether pianists persist in the intensity or the articulation of what they have just heard. The intensity pattern of a musical performance often coordinates with the strong and weak beats in a given meter (Sloboda, 1983). Articulation represents the separation between note events and is measured as the offset time of one event minus the onset time of the next event, so negative values are staccato (separated) and positive values are legato (overlapping).

Computer-generated melodies with either binary or ternary intensity patterns served as stimuli. These prime melodies were produced with either a staccato or a legato articulation across all note events, unrelated to the intensity pattern. The notated musical stimuli, known as target melodies, were metrically ambiguous and contained no bar lines or articulation cues. They could be performed in either binary or ternary meter. The goal of the experiment was to test whether pianists persist in the performance cues that are structurally-related (intensity) or structurally-unrelated (articulation) (Jungers, 2003).

In the experiment, pianists first sight-read two of the notated melodies to assess any bias for performing in one meter or another. Then on each trial, pianists heard a computer-generated prime melody and performed a similar target melody (with a similar number of events and musical structure). The prime melodies were blocked by meter. Pianists were instructed to concentrate on the melodies for a later recognition task.

Pianists performed with a more separated style following the staccato than the legato prime melodies. Thus, pianists persisted in the metrically-unrelated cue of articulation. Pianists did not show a significant difference in intensity on metrically strong and weak events, although there was a trend for more intense events on metrically strong beats. However, pianists did incorporate the meter into their performances. The events that were expected to be more intense (if pianists persisted in the meter from the prime) were instead played with more length. Thus, instead of producing a strong-weak intensity pattern for a target melody that followed a binary prime melody, pianists produced a long-short articulation pattern. Pianists used articulation cues rather than intensity cues to produce a binary or ternary metrical interpretation. This means pianists perceived the meter and persisted in the meter, but they instantiated the meter with different performance cues they those they had heard in the prime melody. Thus, the pianists' performances revealed persistence of metrically-related and metrically-unrelated performance dimensions (Jungers, 2003).

Persistence in speech

When two people engage in conversation, how does the way one person speaks influence the speech of the other? Do conversation partners persist in the speech patterns of one another? One aspect of speech that persists is the syntactic form. Listeners who repeated a sentence they

had heard were more likely to use the same syntax from the sentence when describing a picture (Bock, 1986). For example, when subjects heard and repeated the passive sentence, "The referee was punched by one of the fans," they were more likely to describe a picture with a lightning bolt and church in the passive form as "The church is being struck by lightning" instead of "Lightning is striking the church" (Bock, 1986). Such structural priming lasts beyond one or two sentences (Bock & Griffin, 2000).

There is also evidence that speech rate persists. In one study, participants heard a recording of two male speakers: one at a fast rate and one at a slow rate (Kosslyn & Matt, 1977). Participants then read a paragraph they were told was written by one of the two speakers. The participants read the passage faster when they thought the fast-speaking person wrote it. However, the participants did not perfectly mimic the rate; their production rates were always slower than the rate of the fast speaker (Kosslyn & Matt, 1977).

Further evidence for rate persistence was demonstrated by Jungers, Palmer, and Speer (2002). Participants first read two sentences aloud as a measure of their preferred speech rate. Next, they heard a prime sentence and then read a written target sentence that was matched for number of syllables, lexical stress pattern, and syntactic structure. For example, participants heard, "She read the paper in a hurry" and then read aloud, "He smelled the coffee for a moment" (Jungers et al., 2002). The prime sentences were recorded by a naive female speaker at slow (750 ms or 80 bpm per accent) and fast (375 ms or 160 bpm per accent) rates. As in the music task, participants were instructed to attend carefully to the sentences for a later recognition task. The participants' rates showed an influence of both the prime rate and their preferred speaking rate. Although the speakers and the musicians in the parallel task were both influenced by the prime and preferred performance rates, the speakers were more influenced by the preferred rate and the musicians were more influenced by the prime rate.

Rate is a global prosodic cue that affects the entire utterance. Do speakers also persist in the fine-grained prosodic details of sentences they have heard? Jungers (2003) examined whether prosodic details such as phrase break location and pitch pattern persist in speech. Speech is produced in phrases, often with a pause at a phrase break location. These phrases make the sentence meaning clear and correlate with the sentence structure. The pitch pattern of a sentence can carry meaning, such as a rising pitch at the end of a sentence in English indicating a question. The pitch patterns used in this study occurred at the phrase break locations, but they did not independently add meaning to the sentence. Thus, the phrase break locations were structurally-related while the pitch patterns were structurally-unrelated.

The experimental stimuli were syntactically ambiguous sentences. For example, "Either Brett or Mike and Kay will come to babysit" can be produced with a prosodic phrase break after Brett, which implies that Brett alone or Mike and Kay together will come. This sentence could also be produced with a phrase break after Mike, which implies that Kay will come and one of the two men will also come (Jungers, 2003). A naïve female speaker recorded four versions of each sentence with an early or late phrase break and two different pitch patterns. The written version of each sentence was presented on a computer screen in capital letters without punctuation marks.

The participants first read three sentences aloud to assess their preferred prosodic production. Then participants listened to a prime sentence and produced a target sentence on each trial. Trials were blocked by phrase break location. Participants were instructed to pay careful attention to the sentences for a later recognition task.

Participants used similar phrase breaks in their target sentences as they had heard in the prime sentences. They incorporated the structurally-related phrasing cues into their own utterances. The musicians in the parallel task incorporated both structurally-related (metrical) and structurally-unrelated cues into their performance. Although listeners did not persist in the specific pitch pattern of the prime sentences, this may be partly because the participants did not use the full range of pitch patterns and phrasing. Current research is exploring whether listeners also incorporate pitch from perception to production.

Why persist?

Recent research demonstrates persistence of performance details in both speech and music. This persistence is not limited to tempo. Instead, it includes structurally-related details such as meter and phrasing, as well as structurally-unrelated details such as articulation. What is the advantage of persisting in these performance variations? One possibility is that persistence aids communication in speech and music. By producing similar prosodic patterns, conversation partners may be able to more quickly understand each other. Speakers adjust their utterances to aid listeners. For example, speakers add fillers such as "um" and "uh," that may help to pace the conversation (Clark, 2002). Also, special words and phrases such as "so," "now," "uh," and "um" serve to mark either a new turn or a continuing turn (Fox Tree, 2000). Prosodic persistence may be another way speakers adjust so listeners are prepared for the utterances and can understand easily.

A second possible advantage of persistence is that similar prosodic cues may be easier to produce because they are already primed in the speaker's memory. Syntactic persistence effects in speech have been explained as a type of implicit learning (Bock and Griffin, 2000). Perhaps a similar implicit learning explains persistence of performance cues in music and speech. Exposure to a longer stimulus could lead to a stronger memory representation or to greater implicit learning.

Music and language – common mechanisms?

Does the persistence of performance cues in speech and music stem from common mechanisms? In the study of rate persistence, both musicians and speakers were influenced by their preferred rate as well as the prime rate (Jungers, Palmer, Speer, 2002). However, there are differences in the degree of influence, with the musicians showing greater influence of the prime rate and speakers showing greater influence of their preferred rate. This difference may be due to differing performance expectations. Musicians traditionally perform in an ensemble where the goal is to synchronize with the conductor and fellow performers. Speakers, on the other hand, are more concerned with presenting a clear message and there is little pressure to speak ata particular rate.

The examination of prosodic persistence of structurally-related and structurally-unrelated cues also revealed a distinction between music and language (Jungers, 2003). The musicians showed persistence of metrically-unrelated cues, but they persisted with metrically-related cues to a lesser degree. The speakers persisted in structurally-related cues, but there was little evidence of persistence for structurally-unrelated cues. This difference may be due to the specific task or it may point to a more fundamental difference between the domains.

There is some evidence for a relationship between prosodic and musical patterns. One study examined the perception of two "amusic" subjects who had specific music perception impairments due to brain damage (Patel et al., 1998). The subjects performed a prosodic discrimination task as well as a parallel music discrimination task with stimuli derived from the language task. One subject performed well on both tasks while the other subject performed poorly on both, suggesting a common neural mechanism for interpreting linguistic and musical prosody (Patel et al., 1998). Another study showed that classical compositions by French and English composers differed in rhythmic patterns, paralleling the rhythmic differences between the two languages (Patel & Daniele, 2003). This study suggests linguistic prosody influences musical prosody. This connection between prosody in music and speech is unique, since many aspects of music processing are thought to be specific to music and may be organized modularly (Peretz & Coltheart, 2003).

The literature reviewed here suggests both commonalities and differences in the persistence of performance details across the domains of music and speech. Research continues to explore this persistence effect and the relationship between music and speech.

REFERENCES

Allbritton, D. W., McKoon, G., Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22,* 714–735.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognition Processses, 11,* 17–67.

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18,* 355–387.

Bock, K., Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General, 129,* 177–192.

Bradlow, A. R., Nygaard, L. C., Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics, 61,* 206–219.

Cathcart, E. P., Dawson, S. (1928). Persistence: A characteristic of remembering. *British Journal of Psychology, 18,* 262–275.

Clark, H. H. (2002). Speaking in time. *Speech Communication, 36,* 5–13.

Clynes, M., Walker, J. (1986). Music as time's measure. *Music Perception, 4,* 85–119.

Collier, G. L, Collier, J. L. (1994). An exploration of the use of tempo in jazz. *Music Perception, 11,* 219–242.

Cooper, G., Meyer, L. B. (1960). *The rhythmic structure of music.* Chicago: University of Chicago Press.

Cutler, A., Dahan, D., van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40,* 141–201.

Dowling, W. J., Harwood, D. L. (1986). *Music cognition.* Orlando: Academic Press.

Drake, C. (1993). Perceptual and performed accents in musical sequences. *Bulletin of the Psychonomic Society, 31,* 107–110.

Drake, C., Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10,* 343–378.

Fox Tree, J. E. (2000). Coordinating spontaneous talk. In L. Wheeldon (Ed.) *Aspects of language production* (pp. 375–406). Philadelphia: Psychology Press.

Fox Tree, J. E., Meijer, P. (2000). Untrained speakers' use of prosody in syntactic disambiguation and listeners' interpretations. *Psychological Research, 63*, 1–13.

Henderson, M. T. (1936). *Rhythmic organization in artistic piano performance* (pp.281–305). See CE Seashore.

Jungers, M. K. (2003). *Prosodic persistence in music performance and language production.* Unpublished doctoral dissertation, Ohio State University, Columbus.

Jungers, M. K., Palmer, C. (2000). Episodic memory for music performance. *Abstracts of the Psychonomic Society, 5*, 105.

Jungers, M. K., Palmer, C., Speer, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing, 2*, 21–35.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., Gerken, L. A. (1995). The head-turm preference procedure for testing auditory perception. *Infant Behavior and Development, 18*, 111–116.

Kjelgaard, M. M., Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*, 153–194.

Kosslyn, S. M., Matt, A. M. (1977). If you speak slowly, do people read your prose slowly? Person-particular speech recoding during reading. *Bulletin of the Psychonomic Society, 9*, 250–252.

Large, E. W., Palmer, C. (2001). Perceiving temporal regularity in music. *Cognitive Science, 26*, 1–37.

Large, E. W., Palmer, C., Pollack, J. B. (1995). Reduced memory representations for music. *Cognitive Science, 19*, 53–96.

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa, 7*, 106–122.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics, 5*, 253–263.

Lehiste, I., Olive, J. P., Streeter, L. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America, 60*, 1199–1202.

Lerdahl, F., Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Levitin, D. J., Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics, 58*, 927–935.

Nygaard, L. C., Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60(3)*, 355–376.

Nygaard, L. C., Sommers, M. S., Pisoni, D. B. (1993). Effects of stimulus variability on perception and representation of spoken words in memory. *Perception & Psychophysics, 57*, 989–1001.

Palmer, C. (1996). Anatomy of a performance: sources of musical expression. *Music Perception, 13*, 433–454.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115–138.

Palmer, C., Jungers, M. K., Jusczyk, P. W. (2001). Memory for musical prosody. *Journal of Memory and Language, 45*, 526–545.

Palmer, C., Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 728–741.

Patel, A. D., Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*, B35–B45.

Patel, A. D., Peretz, I., Tramo, M., Labreque, R. (1998). Processing Prosodic and Musical Patterns: A Neuropsychological Investigation. *Brain and Language, 61*, 123–144.

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

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson, J. W. Mullenix (Eds.) *Talker variability in speech processing* (pp. 9–32). San Diego: Academic Press.

Price, P., Ostendorf, M., Shattuck-Hufnagle, S., Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America, 90*, 723–735.

Raffman, D. (1993). *Language, music, and mind.* (Bradford Book, ed.). Cambridge, MA: MIT Press.

Repp, B. H. (1992). Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Cognition, 44*, 241–281.

Seashore, C. E. (Ed.) (1936). *Objective Analysis of Musical Performance* (vol. 4). Iowa City: University of Iowa Press.

Siegel, J. A., Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception and Psychophysics, 21*, 399–407.

Sloboda, J. A. (1983). The communication of music metre in piano performance. *Quarterly Journal of Experimental Psychology, 35*, 377–395.

Sloboda, J. A. (1985). Expressive skill in two pianists: metrical communication in real and simulated performances. *Canadian Journal of Psychology, 39*, 273–293.

Speer, S. R., Crowder, R. G., Thomas, L. M. (1993). Prosodic structure and sentence recognition. *Journal of Memory and Language, 32*, 336–358.

Warren, R. M. (1985). Criterion shift rule and perceptual homeostasis. *Psychological Review, 92*, 574–584.

Wingfield, A., Klein, J. F. (1971). Syntactic structure and acoustic pattern in speech perception. *Perception and Psychophysics, 9*, 23–25.