Language Familiarity, Expectation, and Novice Musical Rhythm Production

John G Neuhoff and Pascale Lidji

Language and Speech published online 12 February 2014
DOI: 10.1177/0023830914520837

Original Article

Language and Speech

Language Familiarity, Expectation, and Novice Musical Rhythm Production

Language and Speech
1–10

Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0023830914520837
las.sagepub.com
SAGE

John G Neuhoff
The College of Wooster, USA

Pascale Lidji
McGill University, Canada

Abstract

The music of expert musicians reflects the speech rhythm of their native language. Here, we examine this effect in amateur and novice musicians. English- and French-speaking participants were both instructed to produce simple "English" and "French" tunes using only two keys on a keyboard. All participants later rated the rhythmic variability of English and French speech samples. The rhythmic variability of the "English" and "French" tunes that were produced reflected the perceived rhythmic variability in English and French speech samples. Yet, the pattern was different for English and French participants and did not correspond to the actual measured speech rhythm variability of the speech samples. Surprise recognition tests two weeks later confirmed that the music–speech relationship remained over time. The results show that the relationship between music and speech rhythm is more widespread than previously thought and that musical rhythm production by amateurs and novices is concordant with their rhythmic expectations in the perception of speech.

Keywords

rhythm, speech, music, native language

Music and language have long been proposed to have common cognitive underpinnings (see Patel, 2007). In many ways the similarities between the two domains have been patently obvious yet empirically elusive. Few doubt that in some ways music can be *like* speech. Philosophers and poets have evidenced similarities between music and speech for thousands of years. Unfortunately, much of this past evidence has been metaphorical or, at best, analogical in nature. However, recent experimental

Corresponding author:
John G Neuhoff, Department of Psychology, The College of Wooster, 930 College Mall, Wooster, OH 44691, USA.
Email: jneuhoff@wooster.edu

work suggests that there is substance to these metaphorical and analogical relationships (for reviews see Besson, Chobert, & Marie, 2011; Kraus & Chandrasekaran, 2010; Patel, 2007). Much of the evidence for similarities between the two domains has focused on pitch (e.g., Deutsch, 1997; Deutsch, Henthorn, & Lapidis, 2011; Deutsch, Henthorn, Marvin, & Xu, 2006; Magne, Schon, & Besson, 2006; Tierney, Dick, Deutsch, & Sereno, 2013). Less work has examined links in rhythm, perhaps because of the relative difficulties in precisely measuring and specifying speech rhythms.

However, one line of work has shown that there may be more than metaphor that ties the two domains together rhythmically. Patel and Daniele (2003) analyzed over 300 pieces of classical instrumental music by French and English composers. They compared stress patterns in the music and language using a normalized Pairwise Variability Index (nPVI), which gives a measure of the average contrast in duration between successive syllables or musical notes. Their results showed that, although there was variability in the rhythmic patterns of individual pieces, the rhythm in the music generally reflected the rhythmic stress patterns, or "prosody", of the native languages of the composers. Even though the acoustic correlates of these categories are still highly debated (see Grabe & Low, 2002, Ramus, Nespor, & Mehler, 1999; White, Mattys, & Wiget, 2012), French has traditionally been characterized as a "syllable-timed language" with less variability in the duration of successive syllables, whereas English, considered a "stress-timed language", has a more variable rhythmic pattern (Abercrombie, 1967; Pike, 1945). Huron and Ollen (2003) showed that this link between composed music and speech rhythms exists across many cultures, replicating the finding with over 7000 pieces of music from 12 different languages. More recent work has even found that musical rhythms produced by musicians with the same native language correspond to the different speech rhythms of their different dialects (McGowan & Levitt, 2011).

As striking as these results are, the samples examined in these studies leave the broader question of the generalizability of the rhythmic relationship between music and speech somewhat unanswered. Classical composers and professional musicians constitute a small and highly specialized population. They have extensive musical training and develop expertise and an elaborate cognitive musical framework or schema that enables them to perceive, organize, and produce music in a way that novices do not (Abe & Hoshino, 1990; Halpern & Bower, 1982; Jordan & Shepard, 1987). Musicians have even been shown to have advantages in processing some low-level aspects of the speech signal (Marie, Kujala, & Besson, 2012; Sadakata & Sekiyama, 2011).

However, making music has generally been a more egalitarian and social endeavor practiced by a much larger proportion of society than a select group of expert musicians. It may be that rhythmic links between music and speech are far more widespread, occurring in amateur and novice musicians as well as experts. For example, recent work has shown that participants with only modest levels of musical experience can correctly classify English and French musical rhythms, even when the two classes are given fictional names (Hannon, 2009).

Here we hypothesized that among amateur and novice musicians, native language would influence both the perception of speech rhythm and the rhythmic variability produced under two different language instruction conditions. Specifically, we predicted that a non-native language of another rhythm class would be perceived as more rhythmically variable than a native language regardless of the actual rhythms inherent in the speech samples. We based this hypothesis on previous work that has shown evidence for "language-specific listening" (Cutler & Otake, 1994; Pallier, Christophe, & Mehler, 1997). This work shows that listeners are adept at perceiving and exploiting the perceptual regularities (e.g., rhythmic pattern) of their own language in order to effectively process speech. However, when listening to an unfamiliar language, they can apply these native language perceptual abilities inappropriately. For example, native English speakers have an expectation for stress-timed speech rhythms that can facilitate native language comprehension and intelligibility (Quene & van Delft, 2010). However, "listening in English" to a syllable-timed French speech sample may produce violations of these rhythmic expectations and produce higher estimates of speech rhythm variability for the unfamiliar language. Moreover, some work has shown culturally specific advantages in processing familiar versus unfamiliar rhythms (Hannon, Soley, & Ullal, 2012; Nazzi, Jusczyk, & Johnson, 2000).

Our second hypothesis was that participants with limited musical expertise would produce simple musical tunes that matched in rhythm their perception of the rhythm of the familiar and unfamiliar languages. We expected participants to produce tunes with greater rhythmic variability when instructed to do so in a non-native language. For example, if English-speaking novice musicians are asked to compose a "French" tune we expect greater variability than if they are instructed to compose in their native English.

These instructions and the open-ended nature of the task leave open the possibility that participants may rely on non-linguistic information to produce their tunes. However, we wanted to avoid strongly invoking language-based contrasts (e.g., providing model exemplars of English and French speech on which participants could base their tunes). In this way we could examine rhythm production without overtly influencing participants to rely on language-based rhythm production. The drawback to this approach is that it leaves open a determination of exactly what mechanisms listeners use to produce novel musical rhythms. The advantage is that it provides a more ecologically valid rhythm production situation. This manipulation brings us closer to conditions used in previous work on classical composers (Huron & Ollen, 2003; Patel & Daniele, 2003). These composers were free to produce any rhythms that they chose. However, we have no idea on what these rhythms are based other than our knowledge of the native language of the composers. In the current work we have not only knowledge of the participants' native language, but also a measure of their perception of speech rhythms in their own and an unfamiliar language.

Testing the production of musical rhythm among non-musicians poses a challenge that we met with a novel methodology. We wanted something that approximated true musical production (e.g., not just finger tapping), but something simple enough that it would not be intimidating. Thus, we asked our native English-speaking and native French-speaking amateurs and novices to use two keys on a computer keyboard to create and record simple "two-note tunes". We chose the two-note-tune method because it provided a relatively simple (somewhat) musical means by which music novices could compose musical rhythms without being intimidated by an unfamiliar musical instrument. Tunes were limited to two notes in an effort to force participants to rely more on rhythmic than melodic characteristics in a later recognition task. Participants could alternate between the two notes as often as they liked and in any rhythm that they desired.

Each participant created two different tunes, one which they were instructed to make a "French" or an "English" song. Two days later the same participants listened to French and English speech samples and rated the perceived variability in rhythmic syllable stress. Finally, two weeks after the initial session, participants were given a surprise recognition task and asked to pick the songs that they had earlier produced from foil songs selected from the rest of the sample.

I Method

1.1 Participants

Native English-speaking participants were 18 female and six male undergraduate students living in the United States who ranged in age from 18 to 22 years. All were native speakers of American English and did not speak French. Some had limited experience playing a musical instrument ($\bar{x}$ years = 5.1, SD = 3.8). None were college music majors or professional musicians. All received course credit for participation.

Native French-speaking participants were 21 female and three male students living in Montreal, Canada, who ranged in age from 18 to 31 years. All were native speakers of French and spoke French as their primary language at home, school, and work. All grew up in a household where French was the primary language and both parents were native speakers of French. All had received some English language instruction beginning at a mean age of 9.7 years (SD = 1.75). Some had limited experience playing a musical instrument ($\bar{x}$ years = 6.4, *SD* = 3.3). None were college music majors or professional musicians. They received a monetary compensation for participation.

1.2 Design and procedure

1.2.1 Musical rhythmic production. Our novice participants found the "two-note-tune" task to be a non-threatening way to produce musical rhythm. They were instructed to create two-note tunes using only two adjacent keys on a computer keyboard. Pressing the keys produced synthesized piano sounds at C4 (261 Hz) and D4 (293 Hz) respectively. Tunes were recorded using a software-based MIDI sequencer without a metronome or any specified tempo. English-speaking participants were instructed to compose two songs, one of which was to be a "French" song and the other was to be a "regular" song. As a conservative measure, we avoided calling the other condition "English" so as not to draw explicit attention to a language-based contrast between the conditions (or suggest a British theme). Similarly, French-speaking participants were instructed to produce two songs, one of which had to be "English" and the other which was to be "regular". It is important to note that we specifically did *not* elaborate on these instructions. We did not want to guide our participants into consciously producing specific "language-based" musical rhythms. Rather we were interested in whether participants would naturally produce language-based rhythms in the absence of elaborate musical schemata. All participants were tested in their native language. Participants were given one minute of rehearsal to familiarize themselves with the task, and any questions that they had were answered. We then recorded their simple two-note tunes of up to 30 seconds in duration. Half of the participants were instructed to produce the tune from their native language first. The nPVI value for each tune was calculated following Patel and Daniele (2003) using the inter-onset interval (IOI) of each note as the unit of analysis:

$$nPVI = \frac{100}{m-1} \times \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{\frac{d_k + d_{k+1}}{2}} \right|$$

where m is the number of intervals in a sequence and d_k is the duration of the kth interval.

1.2.2 Perception of rhythm in speech. We conducted the second part of the experiment two days after the first session to avoid drawing immediate parallels between music and speech rhythms. Participants rated the rhythmic variability of speech samples on a scale of 1–100 after listening to metronomic examples of rhythmic variability where the endpoints of the scale were illustrated by click tracks. A "temporally regular" rhythm was illustrated by isochronous clicks with 500 ms IOIs. A "variable" rhythm was illustrated by a random pattern of clicks with IOIs that ranged from 100 to 900 ms. Participants rated two French samples and two English samples with one male and one female voice in each language. All speech samples were recorded from radio news programs. Speech samples were approximately 11 s in duration and each language was matched for an equal

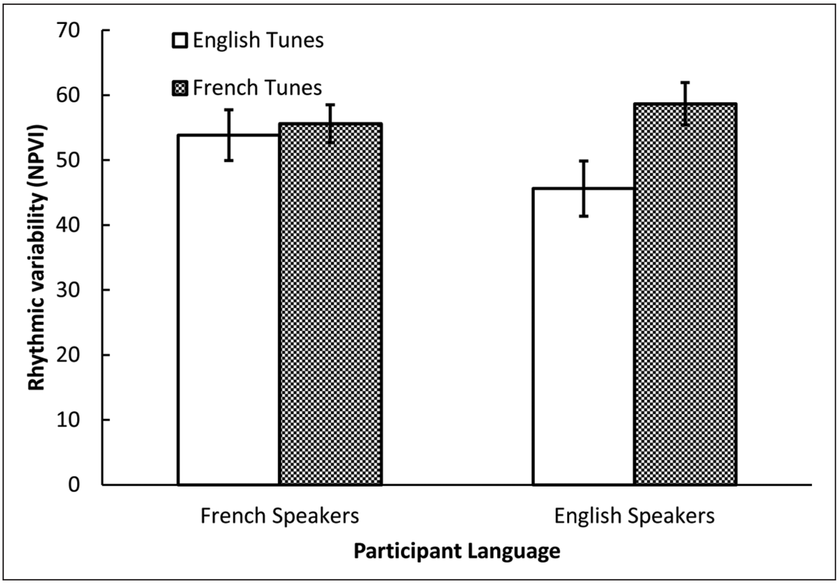

Figure 1. Rhythmic variability produced by French and English participants under instructions to produce either a French or English song. English-speaking participants produce greater rhythmic variability when asked to produce a "French" song. There is no difference between conditions for French-speaking participants. Error bars represent ±1 standard error of mean.

number of syllables. Average nPVI was 52.2 for the English speech samples and 35.5 for the French speech samples.

1.2.3 Recognition task. In surprise recognition tasks two weeks after the initial session, participants were required to identify their tunes from foil tunes in two two-alternative forced choice (2AFC) tasks. All of the participants' tunes served as a target and a foil once in each language condition. For example, Participant 1 heard his or her target tune paired with the tune of Participant 2, used as a foil. Participant 2 heard his or her target tune with Participant 3's tune used as a foil, etc. Foils and targets were always from the same language condition. For the recognition task, all tunes were truncated to the first 10 s to avoid cues to identity based on length. Participants clicked icons on a computer screen to hear the two alternatives in headphones. Half of the participants were tested on the native language tune first.

2 Results

2.1 Musical rhythmic production

The mean nPVI values for the tunes in each condition are shown in Figure 1. We conducted a 2×2 mixed analysis of variance (ANOVA) with the within-subjects factor "Tune Language" (English/French) and the between-subjects factor "Participant Language" (English/French). The main effect for Participant Language was not significant ($F_{(1, 46)} = .47$, $p = .50$). We found a significant main effect for Tune Language ($F_{(1, 46)} = 10.78$, $p = .002$), indicating that French tunes (M = 57.1, SD = 17.0) were more rhythmically variable than English tunes (M = 49.7, SD = 13.6). However, this main effect was driven primarily by a significant interaction between Native Language and Tune

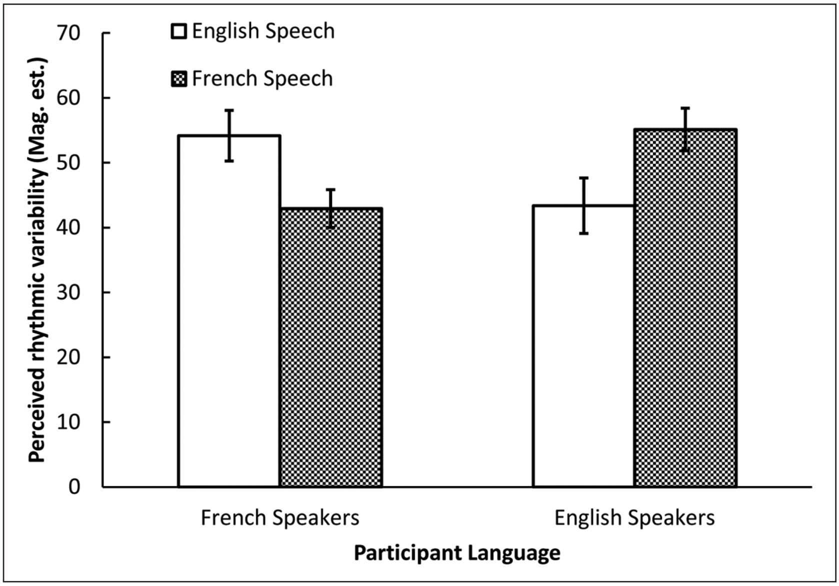

Figure 2. Mean perceived speech rhythm variability of French and English speech tokens by French and English participants. Error bars represent ±1 standard error of mean.

Language ($F_{(1, 4)}$ = 6.26, p = .016). Post hoc comparisons showed that English-speaking participants produced tunes that were significantly more variable in the "French Tune" condition ($t_{(1, 23)}$ = 3.38, p = .003). There was no significant difference in rhythmic variability between the instruction conditions for the French-speaking participants ($t_{(1, 23)}$ = .75, p = .75).

2.2 Perception of rhythm in speech

The mean perceived rhythmic variability for the speech samples in each condition are shown in Figure 2. We conducted a 2×2 mixed ANOVA with the within-subjects factor "Talker Language" (English/French) and the between-subjects factor "Participant Language" (English/French). Main effects for Talker Language ($F_{(1, 46)}$ = .006, p = .94) and Participant Language ($F_{(1, 46)}$ = .021, p = .88) were not significant. However we found a significant interaction between Talker Language and Participant Language ($F_{(1, 43)}$ =10.98, p = .002). English-speaking participants found the French-language samples significantly more variable than the English samples $t_{(1, 23)}$ = 2.89, p = .008). However, for French-speaking participants (who had more exposure to English than our English-speaking participants had to French) the difference in the rhythmic variability of English and French did not reach statistical significance $t_{(1, 23)}$ = 1.99, p = .06 (but was clearly a trend).

2.3 Recognition task

In the delayed recognition tests participants heard a recording of a tune that they had produced (two weeks prior) along with matched-language foil tunes selected from the sample. They were asked to choose which of the two tunes was theirs. Separate recognition tasks were performed for both the French and English tunes. One French-speaking participant failed to return for the recognition task, so the French data were analyzed with an N =23.

English-speaking participants recognized their own two-note tunes at rates significantly above chance levels for both English (83%, $\chi^2_{(1)} = 5.33$, $p < .05$) and French (88%, $\chi^2_{(1)} = 6.75$, $p < .05$) compositions. The difference between recognition of French and English compositions was not significant. French-speaking participants also identified their own two-note tunes at rates significantly above chance levels for both English (91%, $\chi^2_{(1)} = 7.85$, $p < .05$) and French (82%, $\chi^2_{(1)} = 4.89$, $p < .05$) compositions. The difference between recognition of French and English compositions for French speakers was also not significant.

3. Discussion

Our results show that in the absence of a complex cognitive musical schema provided by a lifetime of study and expertise, amateur and novice musicians produce musical rhythms that mirror their perception of the rhythm of a familiar and an unfamiliar language. Native speakers of English and French who had little musical training produced musical rhythms that matched their rhythmic perception of the English and French languages. When asked to produce a French two-note tune, English-speaking participants produced greater rhythmic variability than they did when asked to produce a "regular" tune. This pattern of results mirrors exactly the perception of speech rhythm that the English-speaking participants exhibited when asked to rate the rhythmic variability of French and English speech samples. Despite the greater objective rhythmic variability of the English speech samples, English speakers perceived French as more rhythmically variable and produced "French" and "English" musical rhythms that reflected this linguistic perception. Although our English-speaking sample had little exposure to French speech rhythms, previous work has shown that such participants can readily classify English and French rhythms (Hannon, 2009).

Our French-speaking participants had some familiarity with English, and this is also reflected in our results. French speakers produced English and "regular" tunes that did not differ in their rhythmic variability. Similarly, they did not rate French and English speech samples as being significantly different in rhythm. Thus, for our French-speaking participants, rhythmic production was also largely concordant with their perception of speech rhythms of English and French. However, French-speaking participants showed a non-significant trend to perceive French speech as more regular than English speech. This pattern is opposite that produced by our English-speaking participants and may reflect the greater familiarity that French participants had with French than with English.

There are at least two plausible explanations for the differences in perceived speech rhythm variability across languages. Participants may simply perceive a non-native language as more variable because it is unfamiliar. The well-engrained rhythmic expectancies that are present in one's native language are violated when listening to a language with a different rhythmic pattern. This may produce greater perceived rhythmic variability (Cutler & Otake, 1994; Pallier et al., 1997). Thus, even though English has greater rhythmic variability than French, English may sound less variable to native speakers because they know what to expect rhythmically. Our French-speaking sample had more exposure to English than our English-speaking sample had to French. If exposure to a particular language influences rhythmic expectations and in turn the perception of rhythmic variability, then the pattern of results for our French-speaking participants would be expected.

Another possibility is that when rating the rhythmic variability of speech samples, participants are attending to the semantic message in their native language and therefore spend less attention to the rhythmic variability. Attention to semantic information occurs automatically even when it is not required (Connolly, Stewart, & Phillips, 1990). The inability to attend to the message of the unfamiliar language may allow listeners to spend more of their attention on the rhythmic characteristics of the speech signal. Of course, these two possibilities are not mutually exclusive and may both contribute to our findings. Future work on rating speech variability should tease out the effects of attending to semantics on the perception of rhythmic variability (e.g., with low-pass filtered speech).

The production of these novel tunes appears to be based on a relatively stable rhythmic representation. The surprise recognition test two weeks after producing the tunes showed that despite extremely limited melodic information and no advanced knowledge of a recognition test, both English- and French-speaking participants could identify the tunes that they had composed with high rates of accuracy.

A limitation of the current work is that it is possible that when asked to produce novel tunes, our participants simply played the rhythm to a tune that was familiar to them or followed the rhythm of lyrics to a familiar song. However, this is unlikely for several reasons. Firstly, this would not explain the significant interaction between native language and instruction conditions for the tunes they produced. Moreover, it would be a very difficult task for novices to imagine one familiar tune while playing (and listening to) a melody that did not match the one they were imagining. Finally, when one listens to the tunes produced by these musical novices, it becomes quickly apparent that the organized rhythmic structure of any familiar tune is sorely lacking.

It should be noted that this "linguistic limitation" is also one that should be extended to the previous work with experts (Huron & Ollen, 2003; McGowan & Levitt, 2011; Patel & Daniele, 2003). Although these previous researchers were careful to exclude music with lyrics from their analysis, it is impossible to tell whether any linguistic phrases were in the minds of the composers as they wrote. That the work was composed without lyrics does not necessarily mean that it was not influenced by linguistic phrases. In fact, the direct influence of language could even be more likely in the data on experts. The experts at the very least had more time to contemplate the rhythms that they wrote than our participants, who had to spontaneously perform.

Although it is tempting to compare our results with amateur and novice musicians to the data on experts (Huron & Ollen, 2003; Patel & Daniele, 2003) to examine expertise differences, our methods and data do not permit a stringent direct comparison. The subjects of the research on expert composers were simply composing music. They were not instructed to compose "in a language" as our participants were. Moreover, they were not specifically asked to contrast their native language composition with an unfamiliar language. There may also be important differences between the "pensive composition" of music and spontaneous rhythmic production of a two-note tune. Nonetheless, our results are consistent with this previous work in that we do find evidence of a rhythmic relationship between music and language, even among amateurs and novices.

In a broader sense, our results are consistent with a common foundation for music and speech. Although the evolutionary origins of music and vocal communication are speculative at best, a link between the two has long been proposed (Darwin, 1871; Fenk-Oczlon & Fenk, 2009; Patel, 2006). Previous work shows evidence for this link among experts (Huron & Ollen, 2003; McGowan & Levitt, 2011; Patel & Daniele, 2003). Our results suggest that this phenomenon may extend far beyond the population of expert musicians and be present in participants with little to no musical training.

Funding

This work was supported in part by a grant from the National Science Foundation (award number 0843493).

References

Abe, J., & Hoshino, E. (1990). Schema driven properties in melody cognition: Experiments on final tone extrapolation by music experts. *Psychomusicology*, *9*, 161–172.

Abercrombie, D. (1967). *Elements of general phonetics*. Chicago, IL: Aldine.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*, 94.
Connolly, J. F., Stewart, S. H., & Phillips, N. A. (1990). The effects of processing requirements on neurophysiological responses to spoken sentences. *Brain and Language*, *39*, 302–318.
Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language*, 33, 824–844.
Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London, UK: J. Murray.
Deutsch, D. (1997). The tritone paradox: A link between music and speech. *Current Directions in Psychological Science*, *6*, 174–180.
Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *Journal of the Acoustical Society of America*, *129*, 2245–2252.
Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America*, *119*, 719–722.
Fenk-Oczlon, G., & Fenk, A. (2009). Some parallels between language and music from a cognitive and evolutionary perspective. *Musicae Scientiae*, *13*(2), 201–226.
Grabe, E., & Low, E.L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology VII* (pp. 515–546). The Hague, The Netherlands: Mouton de Gruyter.
Halpern, A. R., & Bower, G. H. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology*, *95*, 31–50.
Hannon, E. E. (2009). Perceiving speech rhythm in music: Listeners classify instrumental songs according to language of origin. *Cognition*, *111*, 403–409.
Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 543–548.
Huron, D., & Ollen, J. (2003). Agogic contrast in French and English themes: Further support for Patel and Daniele (2003). *Music Perception*, *21*, 267–271.
Jordan, D. S., & Shepard, R. N. (1987). Tonal schemas: Evidence obtained by probing distorted musical scales. *Special Issue: The Understanding of Melody and Rhythm*, *41*, 489–504.
Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*, 599–605.
Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*, 199–211.
Marie C., Kujala T., & Besson M. (2012). Musical and linguistic expertise influence pre-attentive and attentive processing of non-speech sounds. *Cortex*, *48*(4), 447–457.
McGowan, R. W., & Levitt, A. G. (2011). A comparison of rhythm in English dialects and music. *Music Perception: An Interdisciplinary Journal*, *28*, 307–314.
Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*, 1–19.
Pallier, C., Christophe, A., & Mehler, J. (1997). Language-specific listening. *Trends in Cognitive Science*, *1*, 129–132.
Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, *24*, 99–103.
Patel, A. D. (2007). *Music, Language, and the Brain*. New York, NY: Oxford University Press.
Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, *87*, B35–B45.
Pike, K. N. (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
Quene, H., & van Delft, L. E. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication*, *52*, 911–918.
Ramus, F., Nespor, M., & Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal. *Cognition*, *72*, 1–28.

Sadakata, M., & Sekiayama, K. (2011). Enhanced perception of various linguistic features by musicians: A cross-linguistic study. *Acta Psychologica*, *138*, 1–10.

Tierney, A., Dick, F., Deutsch, D., & Sereno, M. (2013). Speech versus song: Multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cerebral Cortex*, *23*, 249–254.

White, L., Mattys, S. L., & Wiget, L. (2012). Language categorization by adults is based on sensitivity to durational cues, not rhythm class. *Journal of Memory and Language*, *66*, 665–679.