
Volume 111, Number 3, June 2009
ISSN 0010-0277
pp. 281–422
ELSEVIER
COGNITION

Cognition 111 (2009) 404–410

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Perceiving speech rhythm in music: Listeners classify instrumental songs according to language of origin

Erin E. Hannon *

Department of Psychology, University of Nevada, Las Vegas, 4505 Maryland Parkway, Box 455030, Las Vegas, NV 89514-5030, United States
Harvard University, William James Hall, 33 Kirkland St., Cambridge, MA 02138, United States

A R T I C L E I N F O

Article history:
Received 28 March 2008
Revised 4 March 2009
Accepted 9 March 2009

Keywords:
Music perception
Speech prosody
Musical rhythm
Speech rhythm

A B S T R A C T

Recent evidence suggests that the musical rhythm of a particular culture may parallel the speech rhythm of that culture's language (Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*, B35–B45). The present experiments aimed to determine whether listeners actually perceive such rhythmic differences in a purely musical context (i.e., in instrumental music without words). In Experiment 1a, listeners successfully classified instrumental renditions of French and English songs having highly contrastive rhythmic differences. Experiment 1b replicated this result with the same songs containing rhythmic information only. In Experiments 2a and 2b, listeners successfully classified original and rhythm-only stimuli when language-specific rhythmic differences were less contrastive but more representative of differences found in actual music and speech. These findings indicate that listeners can use rhythmic similarities and differences to classify songs originally composed in two languages having contrasting rhythmic prosody.

1. Introduction

Music and speech are both complex acoustic signals that require listeners to anticipate, segment, and interpret elements that unfold over time. Although dissimilar in content and function, increasing evidence suggests that some overlapping perceptual and neural mechanisms underlie music and language processing (Alcock, Passingham, Watkins, & Vargha-Khadem, 2000; Maess, Koelsch, Gunter, & Friederici, 2001; Magne, Schön, & Besson, 2006; Patel, Foxton, & Griffiths, 2005; Wong, Skoe, Russo, Dees, & Kraus, 2007; Patel, Wong, Foxton, Lochy, & Peretz, 2008). For example, elements of speech prosody, such as rhythm and pitch patterning, may engage the same perceptual processes that are needed for music.

Durational patterning is fundamental to the rhythmic structure of both music and speech. The first four notes of Beethoven's Fifth Symphony have a distinctive *short–short–short–long* rhythm, regardless of each note's absolute duration or the speed of performance. Infants' and adults' perception and identification of individual rhythms is based on the relative pattern of long and short durations (Chang & Trehub, 1977; Clarke, 1999; Demany, McKenzie, & Vurpillot, 1977; Lewkowicz, 2003; Trehub & Thorpe, 1989), and cross-cultural differences in the complexity of durational patterns give rise to culture-specific perceptual biases that emerge during infancy (Hannon & Trehub, 2005a, 2005b). Speech also contains distinctive rhythmic properties that vary across cultures. Linguists have intuitively used rhythmic classifications such as "stress-timed" to describe languages like English and Dutch that contain highly varied syllable structure, and "syllable-timed" for languages like French and Spanish containing less varied syllable structure (Grabe & Low, 2002). One potential acoustic basis for rhythmic classes in speech concerns vowel reduction, which is very pronounced in stress-timed languages but weak or absent in syllable-timed languages. Vowel reduction leads to greater variability (durational contrast) between adjacent vowel durations in sentences

* Address: Department of Psychology, University of Nevada, Las Vegas, 4505 Maryland Parkway, Box 455030, Las Vegas, NV 89514-5030, United States. Tel.: +1 702 895 4687; fax: +1 702 895 0195.
E-mail address: erin.hannon@unlv.edu.

0010-0277/$ - see front matter
doi:10.1016/j.cognition.2009.03.003

of stress-timed languages, and lower durational contrast in syllable-timed languages (Grabe & Low, 2002; Ramus, Nespor, & Mehler, 1999; White & Mattys, 2007).

Interest in potential speech-music parallels has been bolstered by recent findings that the rhythmic properties of a particular language are mirrored in the music of that culture (Bispham, 2006; Trehub & Hannon, 2006). Patel and Daniele (2003) adapted a measure of durational contrast, the normalized Pairwise Variability Index[1] (nPVI) to examine the patterning of successive note durations in instrumental music from England and France. They discovered that patterns of note duration in music closely parallel patterns of vowel duration in speech, with successive notes in English music having contrasting long and short durations (i.e., higher nPVI values) and successive notes in French music having more uniform durations (i.e., lower nPVI values). This pattern was replicated with classical music from numerous cultures and eras (Huron & Ollen, 2003) and popular music containing lyrics (Sadakata, Desain, Honing, Patel, & Iversen, 2004).

Parallel rhythmic structures in music and speech might implicate overlapping musical and linguistic representations. Exposure to rhythmic structures begins *in utero* (Abrams, Gerhardt, Huang, Peters, & Langford, 2000), which presumably underlies newborns' preferences for their native language (Bahrick & Pickens 1988; Mehler et al., 1988; Moon, Cooper, & Fifer, 1993) and their ability to discriminate unfamiliar languages that come from contrasting rhythmic classes (Nazzi, Bertoncini, & Mehler, 1998; Ramus, Hauser, Miller, Morris, & Mehler, 2000). Infants' early representations of both speech and music are strongly influenced by statistical properties of input, such as the frequency with which certain events or sequences of events occur (Hannon & Johnson, 2005; Hannon & Trehub, 2005b; Jusczyk, 2002; Maye, Werker, & Gerken, 2002; McMullen & Saffran, 2004). If overlapping culture-specific rhythmic structures occur in both music and language, two structures that are prominent in the auditory environment of young infants (Trehub & Trainor, 1998), it is possible that such input leads to sustained overlapping cognitive representations of music and speech rhythm.

An alternative possibility is that similarities between speech and music rhythm arise indirectly, as historical artifact. For example, because many musical compositions originate in song, the structure of music may inevitably reflect some of the prosodic features of the language to which it was originally set. Thus, language-based rhythms could pervade music but mean very little to the listener in a musical context. Although listeners perceive and respond to various rhythmic structures in music, such as rhythmic motifs and meter (Eerola, Järvinen, Louhivuori, & Toiviainen, 2001; Hannon, Snyder, Eerola, & Krumhansl, 2004; Snyder & Krumhansl, 2001), no evidence to date suggests that listeners perceive the degree to which successive note durations in music are characterized by variability or uniformity. The goal of the present study was to determine whether listeners perceive language-specific rhythmic properties in a musical context.

2. Experiment 1

Rhythmic variability in speech is perceptually salient to the casual listener, as shown by the finding that subjects can classify synthesized speech on the basis of rhythmic class and amount of durational contrast (Ramus, Dupoux, & Mehler, 2003; Ramus & Mehler, 1999). Using a similar paradigm, the present experiments trained listeners to classify a set of French and English instrumental folk songs according to language of origin. Following training, listeners attempted to classify novel French and English songs in the absence of feedback. In one experiment (Experiment 1a), listeners were trained and tested on original folk songs, whereas in a second experiment (Experiment 1b) they were trained and tested on monotone versions of those songs that contained only rhythmic information.

2.1. Method

2.1.1. Participants

Forty-eight adults (25 female, 23 male; ages 18–39 years) participated for course credit or $10 payment. Formal music training among participants ranged from 0 to 19 years ($M = 5.8$). Most participants were native English speakers, but six individuals were born in Vietnam, Russia, Croatia, or Chile and learned English before the age of 5.

2.1.2. Materials

Stimuli were selected by first analyzing a large corpus of French and English songs from anthologies of children's music in English and French. As in prior studies, nPVI[1] for each song was calculated by setting the quarter-note value to 1 and coding all other note durations as multiples or fractions of the quarter-note (Huron & Ollen, 2003; Patel & Daniele, 2003; Patel et al., 2006). Pairwise comparisons included all consecutive notes, excluding those separated by rests. As expected, nPVI was higher for the 75 English songs ($M = 43$) than for the 75 French songs ($M = 33.5$). To maximize the salience of culture-specific rhythmic variation, we selected a subset of 24 songs (12 English and 12 French) having highly contrasting nPVI values (English $M = 60$, French $M = 14$) (see Appendix for a complete list of songs).

For Experiment 1a, instrumental versions of each song were created by entering songs directly from notation into a MIDI sequencer (Digital Performer 4.0), and converted to AIFF format using a Reason 4 sampler *Flute Ensemble* instrument.[2] Each song was thus a simple flute melody con-

[1] The variable nPVI provides an index of durational contrast as shown in the following equation:

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

where m is the number of elements in a sequence and d_k is the duration of the kth element. The absolute difference is calculated between each successive pair of durations, normalized by the mean duration of the pair (Patel, Iversen, & Rosenberg, 2006). Values of nPVI range from 0 to 200, with 200 reflecting the greatest amount of durational contrast.

[2] Examples of English and French songs from Experiments 1 and 2 (Original and Rhythm Only) can be heard at http://faculty.unlv.edu/ehannon/SongStim09.html.

taining no words. Quarter-note durations were set to 600 ms or 100 BPM (beats per minute), which resulted in comparable tempos for English (Number of Notes $M = 47.6$, $SD = 17.8$; Total Duration $M = 20.53$ s, $SD = 8.8$ s) and French songs (Number of Notes $M = 48.6$, $SD = 12.6$; Total Duration $M = 17.7$ s, $SD = 5.1$ s). Pitch values, taken directly from musical notation, were similar for French and English songs (pitch range = 19 semitones), although English songs had a lower mean frequency ($M = 439.6$ Hz) than French songs ($M = 522.8$ Hz).

While rhythmic differences were of primary interest, other non-rhythmic differences might also differentiate French from English songs, such as familiarity, mean frequency, implied harmony, frequency of individual notes, or melodic variability (see Patel et al. (2006)). Experiment 1b therefore drastically limited the cues available for successful performance by presenting the same songs as monotone rhythms. Songs from Experiment 1a were therefore modified in the following fashion. Before being converted to AIFF format, each note was changed to middle C (262 Hz), which preserved rhythmic structure but eliminated pitch variation.

2.1.3. Apparatus and procedure

Participants were tested individually at a computer and presented with instructions and stimuli over headphones by means of Psyscope software (Cohen, MacWhinney, Flatt, & Provost, 1993). The procedure, adapted from Ramus and Mehler (1999), contained a training phase followed by a generalization phase. Songs were randomly divided into training and generalization sets of 12 songs (6 French and 6 English). The training phase began with a sample song from each language accompanied by a fictional label, *Navi* or *Latu*.[3] Participants were then asked to classify all 12 training songs, randomly ordered, as either Navi or Latu. After responding via button press, participants received visual feedback about the language of origin (i.e., "this song was from the Navi language"). Following each trial, participants rated the familiarity of the song on a scale of 1 (very unfamiliar) to 10 (very familiar). All participants completed two blocks of training. The generalization block then presented 12 novel songs (6 from each language), which were presented three times each and ordered randomly with the constraint that no song was repeated consecutively. On each trial, participants classified the song as Navi or Latu and then rated familiarity. No feedback was given during generalization. Half of the participants were trained and tested on the original songs (Experiment 1a) and half on the monotone rhythms (Experiment 1b). In Experiment 1b, training and generalization songs were randomly counterbalanced across subjects so that generalization-block responses could be obtained for all 24 songs.

Following the experiment, each participant completed a questionnaire assessing linguistic background, music training, and any history of hearing impairment.

[3] The fictional labels *Navi* and *Latu* were used instead of actual labels (English and French) to discourage familiarity-based categorization strategies. In other words, we wanted to prevent our English-speaking subjects from assuming they should categorize familiar songs as English and unfamiliar songs as French.

2.2. Results and discussion

2.2.1. Experiment 1a

Table 1 summarizes performance across training and generalization blocks (chance is 50%). Participants learned to classify songs during training, presumably because of explicit feedback. Of most interest was whether participants could then classify a novel set of songs in the absence of feedback. A one-sample *t*-test (two tailed) revealed that accuracy for the generalization block in Experiment 1a was significantly above chance, $t(23) = 9.683$, $p < .001$ (see Fig. 1), supporting the notion that participants classify songs according to language-based rhythmic properties.

It is nevertheless possible that other factors also contributed to performance, such as prior exposure to French or other syllable-timed languages, or prior exposure to particular songs used in the experiment. There was a small but non-significant difference between familiarity ratings for English ($M = 2.04$, $SD = 0.54$) and French songs ($M = 1.59$, $SD = 0.38$), $t = 1.636$, $p = .13$, indicating that English songs were more familiar than French songs. Thus, in principle listeners could have adopted a strategy of grouping familiar songs together. We therefore used correlation and regression to examine which variables predicted classification performance. No significant correlations were found between individual accuracy scores and age, years of formal music training, or years of exposure to a syllable-timed language such as French or Spanish (as a second language) ($p > .27$), so these variables were excluded from further analysis.

To determine which variables were most predictive of listener responses, two variables measuring rhythm (nPVI) and familiarity were regressed onto the averaged response (i.e., tendency to classify a song into the English group) for each of the 12 songs used during the generalization block, yielding a high prediction level, $R^2(2, 9) = .76$, $p < .01$. Table 2 presents simple correlations for each variable. Multiple regressions were then conducted to determine how removal of a given variable would affect the model's ability to predict English responses. Thus, R^2 Change (Table 2) for a given variable indicates the amount by which the predictive strength of the model containing both variables drops when that variable is removed from the regression, and thus reflects the unique contribution of that variable (Darlington, 1990). As shown in Table 2, both nPVI and familiarity were positively and significantly correlated with listeners' tendency to classify individual songs as English, and both variables contributed uniquely to the model. The unique contributions of nPVI and familiarity were similar, indicating that while language-specific rhythmic properties influenced performance, listeners also tended to respond "Navi" (English) whenever songs were more familiar.

2.2.2. Experiment 1b

The use of monotone rhythms in Experiment 1b provided an important means of measuring classification performance while minimizing the contribution of non-rhythmic information. Inspection of Table 1 suggests that performance in Experiment 1b was comparable to Experiment 1a. Accuracy for the generalization block was

Table 1
Mean accuracy scores (percent correct) across Experiments 1–2 for training and generalization blocks. Asterisks indicate performance that was significantly above chance (50%).

Songs	Training A	Training B	Generalization
Experiment 1a (Original melodies)	76.4**	91.3**	77.3**
Experiment 1b (Rhythm-only)	79.2**	86.8**	77.1**
Experiment 2a (Original melodies)	69.1**	71.7**	68.4**
Experiment 2b (Rhythm-only)	58.3*	74.6**	70.5**

* $p = .05$.
** $p < .001$.

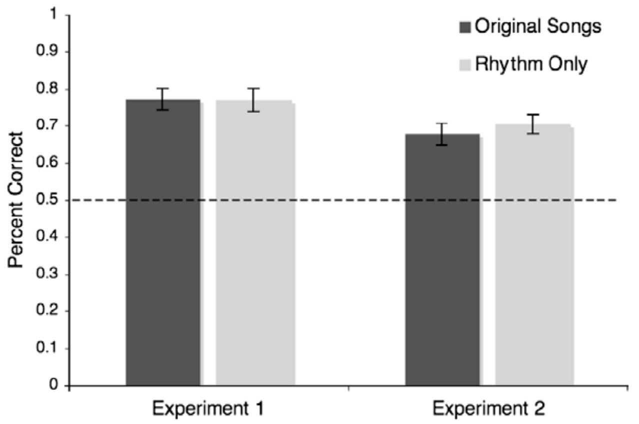

Fig. 1. Mean percent correct (error bars indicate *se*) for Experiments 1a and 1b (songs from French and English characterized by large nPVI differences, with and without pitch information) and Experiments 2a and 2b (songs from French and English characterized by moderate nPVI differences, with and without pitch information). The dotted line indicates chance performance.

Table 2
Simple (r) and R^2 change for variables rhythm (nPVI) and familiarity in Experiment 1.

	Variable	r	R^2 change
Original melodies (Exp. 1a)	nPVI	.77**	.17*
	Familiarity	.65*	.16*
Rhythm-only (Exp. 1b)	nPVI	.86**	.74**
	Familiarity	−.10	.01

* $p < .05$.
** $p < .01$.

significantly above chance, $t(23) = 8.807$, $p < .001$ (see Fig. 1). There were no correlations between accuracy and age, formal music training, or exposure to a syllable-timed language ($p > .16$).

The absence of pitch information degraded the familiarity of songs, minimizing differences between familiarity ratings for English ($M = 1.82$, $SD = 0.48$) and French songs ($M = 1.845$, $SD = 0.45$), $t = -.13$, $p = .90$). A regression analysis with nPVI and familiarity revealed a high prediction level, $R^2(2,\ 21) = .75$, $p < .001$, with only nPVI predicting responses and contributing uniquely to the model (Table 2). Likewise, familiarity was not correlated with classification responses.

In summary, because rhythmic differences provided the only cue to group membership in Experiment 1b, the combined results of Experiments 1a and 1b strongly suggest that adults classified songs based on rhythmic properties.

3. Experiment 2

Experiments 1a and 1b showed that rhythm alone was sufficient for accurate classification of French and English songs having highly contrastive rhythms. This finding could reflect a natural tendency for listeners to abstract language-specific rhythmic properties in both linguistic and musical input. However, the mean difference in nPVI values for French and English songs in Experiment 1 was far greater than differences observed in any musical or speech corpus. It is therefore possible that listeners only noticed these rhythmic properties because they were unnaturally salient, and that less extreme language-specific rhythmic differences would normally go unnoticed in a musical context. Experiment 2 aimed to enhance ecological validity by replicating Experiment 1 using songs more representative of the larger corpus.

3.1. Method

3.1.1. Participants

Forty-eight adults (31 female, 17 male; ages 18–44 years) participated for course credit. The amount of formal music training among participants ranged from 0 to 11 years ($M = 2.9$). Most participants were native English speakers, but nine individuals were born in Vietnam, The Philippines, China, Japan, or Mexico and learned English before the age of 5.

3.1.2. Materials

A subset of 24 songs was selected from the original musical corpus containing nPVI values falling in the middle of the range for each language. Thus, mean nPVI values in this subset were comparable to the mean values of the larger corpus (English $M = 43.3$, French $M = 31.6$) (see Appendix for song list).

Instrumental versions of each song were created as in Experiment 1.[2] Tempo was set at 100 BPM and was thus comparable for English (Number of Notes $M = 48.2$, $SD = 14.9$; Total Duration $M = 24.2$ s, $SD = 7.35$ s) and French songs (Number of Notes $M = 51.7$, $SD = 9.3$; Total Duration $M = 21.4$s, $SD = 4.35$ s). Pitch values were also similar for French and English songs (pitch range = 10 semitones), although French songs had a higher mean frequency ($M = 745$ Hz) than English songs ($M = 680$ Hz). Songs in Experiment 2a contained original pitch and rhythmic patterning whereas songs in Experiment 2b contained only rhythmic information.

3.1.3. Apparatus and procedure

The apparatus and procedure were identical to Experiment 1, except that training and generalization songs were randomly counterbalanced across subjects in both Experiments 2a and 2b.

3.2. Results and discussion

3.2.1. Experiment 2a

As shown in Table 1, performance was significantly above chance levels in the generalization block, $t(23) = 6.203$, $p < .001$. As in Experiment 1a, English songs were rated as more familiar ($M = 5.36$, $SD = 1.06$) than French songs ($M = 4.85$, $SD = 0.69$), $t = 1.385$, $p = .18$. Despite this, familiarity ratings did not correlate with listeners' tendency to classify individual songs as English (see Table 3). Multiple regression revealed a moderate prediction level, $R^2(2, 21) = .466$, $p < .01$, with nPVI but not familiarity contributing uniquely to the model (Table 3). No significant correlations were found between individual accuracy scores and age, years of formal music training, or years of exposure to a syllable-timed language such as French or Spanish ($p > .16$).

3.2.2. Experiment 2b

Performance in Experiment 2b was comparable to Experiment 2a, except that performance during training blocks was initially lower when only rhythmic information was available (Training A). Nevertheless, accuracy in the generalization block was significantly above chance, $t(23) = 7.922$, $p < .001$ (see Fig. 1). There were no correlations between accuracy and age, formal music training, or exposure to a syllable-timed language ($p > .17$).

Removal of pitch variation actually reversed the pattern of familiarity ratings for English ($M = 5.33$, $SD = 0.27$) and French songs ($M = 5.5$, $SD = 0.24$), but this difference was not significant, $t = -1.65$, $p > .11$. A regression analysis with nPVI and familiarity predicted responses, $R^2(2, 21) = .26$, $p < .05$, but only nPVI contributed uniquely to the model (Table 3). There was no significant correlation between familiarity and tendency to classify the song as English.

The positive correlations between nPVI and classification imply that performance accuracy should increase with degree of rhythmic contrast. Thus, lower accuracy would be expected in Experiment 2 when language-specific differences were subtle, relative to Experiment 1 when differences were extreme. Moreover, if rhythm truly drives listeners' classification of songs, accuracy should not differ for conditions in which both pitch and rhythm varied and conditions where only rhythm varied. Both of these predictions were supported by a 2 (song corpus) × 2 (presence or absence of pitch variation) between-subjects ANOVA across all four experiments, which revealed a main effect of song corpus, $F(1,92) = 7.199$, $p < .01$, but no main effect of pitch variation and no interaction. Thus, the songs in Experiment 1 yielded higher accuracy than did songs in Experiment 2, regardless of the presence or absence of melodic information. Despite these differences, the above-chance performance across all four experiments provides strong support for the notion that listeners do perceive language-specific rhythmic structures in an exclusively musical context, and that they do so even when such differences are subtle.

Table 3
Simple (r) and R^2 change for variables rhythm (nPVI) and familiarity in Experiment 2.

	Variable	r	R^2 change
Original melodies (Exp. 2a)	nPVI	.68**	.43**
	Familiarity	.19	.01
Rhythm-only (Exp. 2b)	nPVI	.44*	.18*
	Familiarity	−.28	.07

* $p < .05$.
** $p < .01$.

4. General discussion

The distinctive rhythmic properties of various languages are salient to listeners, but to date no studies have examined whether comparable rhythmic patterns are perceived in music, despite recent evidence of rhythmic parallels between specific languages and the corresponding culture's music (Patel & Daniele, 2003; Huron & Ollen, 2003). The present results indicate that listeners do in fact perceive language-specific rhythms in musical contexts and can use this information to classify purely instrumental sequences (i.e., music without lyrics). They can do this when language-specific differences are exaggerated and when they are subtle.

Formal musical training was uncorrelated with individual performance, which might be surprising in light of known effects of formal music training on numerous auditory tasks (Besson & Faita, 1995; Koelsch, Schröger, & Tervaniemi, 1999). However, increasing evidence suggests that passive exposure to structures of a particular genre or culture powerfully influences music processing, especially in rhythm tasks (Bigand & Poulin-Charronnat, 2006; Dalla Bella & Peretz, 2005; Hannon & Trehub, 2005b; Honing & Ladinig, 2009). Thus, listeners may acquire sensitivity to language-specific rhythms simply by listening to the speech and music of their culture.

The present findings may have implications for current controversies over the extent to which representations of music and speech are domain-general or domain-specific (Peretz & Coltheart, 2003; Patel, 2003; Trehub & Hannon, 2006) – namely, a common mechanism might underlie rhythm perception in both domains. The present experiments show that listeners are sensitive to language-specific rhythms in an exclusively musical context, which suggests that parallels between speech and music rhythm are not merely historical artifact but are experienced by listeners. These results do not, however, unambiguously point towards shared, overlapping representations of music and speech rhythm. For example, because language-specific rhythms are present in the music of a given culture, an individual listener's representation of music rhythm could be shaped exclusively by musical and not linguistic experience. Separate language- and music-specific representations could thus operate separately but in parallel. If rhythmic representations are truly interactive, experience in one domain should influence perception in the other, and future research would be necessary to show that listeners can, for example, classify rhythmic structures across domains. The present study nevertheless lays an important foundation for such future work by establishing that language-specific rhythmic properties not only exist in purely instrumental music, but that they are readily perceived by listeners.

Acknowledgements

I gratefully acknowledge Rachel Levine, Ling Pan, and Yoko Kobayashi for their assistance in analyzing folk songs and collecting data. I thank Kimberley Barchard for her assistance in data analysis.

Appendix

English songs	
1. *Animal Fair* (Leonard, 1998)	(Exp. 1 and 2)
2. *Any Dream Will Do* (Leonard, 1999)	(Exp. 2)
3. *The Bear Went Over the Mountain* (Leonard, 1998)	(Exp. 1 and 2)
4. *Be Kind to Your Web-Footed Friends* (Leonard, 1999)	(Exp. 2)
5. *By'm Bye* (Seeger, 1948)	(Exp. 1 and 2)
6. *Did You Go to the Barney?* (Seeger, 1948)	(Exp. 1)
7. *Fire Down Below* (Seeger, 1948)	(Exp. 1)
8. *Hakuna Matata* (Leonard, 1999)	(Exp. 1)
9. *John Henry* (Seeger, 1948)	(Exp. 2)
10. *Oh, Oh the Sunshine!* (Seeger, 1948)	(Exp. 2)
11. *Rain, Come Wet Me* (Seeger, 1948)	(Exp. 1)
12. *Rain or Shine* (Seeger, 1948)	(Exp. 1)
13. *Rumbly in My Tumbly* (Leonard, 1999)	(Exp. 2)
14. *The Thing* (Leonard, 1998)	(Exp. 1 and 2)
15. *The Wind Blows East* (Seeger, 1948)	(Exp. 1)
16. *This Lady She Wears a Dark Green Shawl* (Seeger, 1948)	(Exp. 2)
17. *Three Little Fishies* (Leonard, 1999)	(Exp. 1)
18. *Tie Me Kangaroo Down Sport* (Leonard, 1998)	(Exp. 1 and 2)
19. *You've Got a Friend in Me* (Leonard, 1999)	(Exp. 2)
French songs	
1. *La Bonne Aventure* (Fassio, 1932)	(Exp. 2)
2. *Chapeau de Paille* (momes.net)c	(Exp. 2)
3. *Charmante Gabrielle* (Byrd, 1903)	(Exp. 1)
4. *Fais Dodo, Colas* (Fassio, 1932)	(Exp. 2)
5. *Gentile Coqu'licot* (Poire, 1962)	(Exp. 2)
6. *Il Etait Un P'tit Homme* (Poire, 1962)	(Exp. 1)
7. *Il Etait Une Dame Tartine* (Poire, 1962)	(Exp. 2)
8. *J'ai Du Bon Tabac* (Fassio, 1932)	(Exp. 2)
9. *La Fontaine* (Byrd, 1903)	(Exp. 1)
10. *Le Furet du Bois Joli* (Fassio, 1932)	(Exp. 1)
11. *La Mer' Michel* (Fassio, 1932)	(Exp. 1 and 2)
12. *La Petit Nigaud* (Poire, 1962)	(Exp. 1)
13. *Maman, Les P'tits Bateaux* (Poire, 1962)	(Exp. 1)
14. *Marie, Trempe Ton Pain* (Poire, 1962)	(Exp. 2)

Appendix (*continued*)

15. *Monsieur de la Palisse* (Fassio, 1932)	(Exp. 1 and 2)
16. *Ne Pleure pas Jeanette* (momes.net)	(Exp. 2)
17. *Nous Etions Dix Filles A Marier* (Poire, 1962)	(Exp. 1)
18. *Petit Papa* (Fassio, 1932)	(Exp. 2)
19. *Semons la Salade* (Davison & Surette, 1922)	(Exp. 1)
20. *Sur Le Pont d'Avignon* (Poire, 1962)	(Exp. 2)
21. *Ramene tes Moutons* (Davison and Surette, 1922)	(Exp. 1)
22. *Savez-vous Planter Les Choux?* (Fassio, 1932)	(Exp. 1)

References

Abrams, R. M., Gerhardt, K. J., Huang, X., Peters, A. J. M., & Langford, R. G. (2000). Musical experiences of the unborn baby. *Journal of Sound and Vibration, 231*, 253–258.

Alcock, K. J., Passingham, R. E., Watkins, K., & Vargha-Khadem, F. (2000). Pitch and timing abilities in inherited speech and language impairment. *Brain and Language, 75*, 34–46.

Bahrick, L. E., & Pickens, J. N. (1988). Classification of bimodal English and Spanish passages by infants. *Infant Behavior and Development, 11*, 277–296.

Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1278–1296.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*, 100–130.

Bispham, J. (2006). Rhythm in music: What is it? Who has it? And why? *Music Perception, 24*, 125–134.

Byrd, J. (1903). *Folk songs and other songs for children*. Boston: Radcliffe-Whitehead.

Chang, H., & Trehub, S. E. (1977). Infants' perception of temporal grouping in auditory patterns. *Child Development, 48*, 1666–1670.

Clarke, E. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (pp. 473–500). New York: Academic Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, and Computers, 25*, 257–271.

Dalla Bella, S., & Peretz, I. (2005). Differentiation of classical music requires little learning but rhythm. *Cognition, 96*, B65–B78.

Darlington, R. B. (1990). *Regression and linear models*. New York: McGraw-Hill.

Davison, A. T., & Surette, T. W. (1922). *140 folk tunes for school and home*. Boston: Schirmer Music Co..

Demany, L., McKenzie, B., & Vurpillot, E. (1977). Rhythm perception in early infancy. *Nature, 266*, 718–719.

Eerola, T., Järvinen, J., Louhivuori, J., & Toiviainen, P. (2001). Statistical features and perceived similarity of folk melodies. *Music Perception, 18*, 275–296.

Fassio, A. (1932). *French folk songs: Nursery rhymes and children rounds*. New York: Marks Music.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm classhypothesis. In C. Gussenhoven & N. Warner (Eds.). *Laboratory phonology* (Vol. 7, pp. 515–546). Berlin: Mouton de Gruyter.

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology, 50*, 354–377.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 956–974.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16*, 48–55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences USA, 102*, 12639–12643.

Honing, H., & Ladinig, O. (2009). Exposure influences expressive timing judgments in music. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 281–288.

Huron, D., & Ollen, J. (2003). Agogic contrast in French and English themes: Further support for Patel and Daniele (2003). *Music Perception, 21*, 267–271.

Jusczyk, P. W. (2002). How infants adapt speech-processing capacities to native-language structure. *Current Directions in Psychological Science, 11*, 15–18.

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive processing in musicians. *NeuroReport, 10*, 1309–1310.

Leonard, H. (1998). *The really big book of children's songs*. Milwaukee: H. Leonard.

Leonard, H. (1999). *The mighty big book of children's songs*. H. Leonard: Milwaukee.

Lewkowicz, D. J. (2003). Learning and discrimination of audiovisual events in human infants: The hierarchical relation between intersensory temporal synchrony and rhythmic pattern cues. *Developmental Psychology, 39*, 795–804.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4*, 540–545.

Magne, C., Schön, 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.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101–B111.

McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception, 21*, 1–23.

Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*, 143–178.

Momes.net, premiers pas sur internet (n.d.). Comptines, chansons, et poesies. Available from <http://www.momes.net/comptines/comptines-chansons.html> (Retrieved November 2008).

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development, 16*, 495–500.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 756–766.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*, 674–681.

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

Patel, A. D., Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception, 25*, 357–368.

Patel, A. D., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition, 59*, 310–333.

Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *Journal of the Acoustical Society of America, 119*, 3034–3047.

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

Poire, H. (1962). *Mon premier livre de chansons*. Larousse: Paris.

Ramus, F., Dupoux, E., & Mehler, J. (2003). The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 337–342). Barcelona, Spain.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science, 288*, 349–351.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America, 105*, 512–521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*, 265–292.

Sadakata, M., Desain, P., Honing, H., Patel, A. D., & Iversen, J. R. (2004). A cross-cultural study of the rhythm in English and Japanese popular music. In *Proceedings of the international symposium on musical acoustics* (pp. 41–44). Nara, Japan.

Seeger, R. C. (1948). *American folk songs for children*. New York: Doubleday.

Snyder, J. S., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception, 18*, 455–489.

Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*, 73–99.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology, 43*, 217–229.

Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: Lullabies and play songs. *Advances in Infancy Research, 12*, 43–77.

White, L., & Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35*, 501–522.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420–422.