CHAPTER I.

Infants Use Meter to Categorize Rhythms and Melodies: Implications for Musical Structure Learning*

1.1. Introduction

Infants are confronted with rapidly changing, complex auditory patterns such as music and speech from before the time they are born. A question of great interest is how infants learn to organize, parse, and interpret these patterns. Increasing behavioral and computational evidence suggests that infants can use distributional properties of input, such as frequency of occurrence, co-occurrence of syllables or units, and the alternation of strong and weak units, to learn about structure in language and music (Christiansen, Allen, & Seidenberg, 1998; Mattys & Jusczyk, 2001; Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996; Thiessen & Saffran, 2003). The term "distributional" refers to the frequency with which elements or combinations of elements occur in any type of input. Rhythmic information might be particularly important in guiding how infants perceive distributional information in auditory patterns. For example, newborn infants appear to discriminate native from non-native languages on the basis of rhythmic structure (Nazzi, Bertoncini, & Mehler, 1998; Nazzi, Jusczyk, & Johnson, 2000), an ability that may facilitate subsequent language learning (Curtin, Mintz, & Christiansen, in press; Cutler, 1994). Rhythmic structure might also play a fundamental role in infants' early musical experiences and music learning. The present work asks whether infants can use the distributional information in music to infer its underlying temporal organization, and whether this temporal organization might function as a framework for learning other complex structures in music.

* 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.

Two interacting aspects of temporal structure characterize musical rhythm: *rhythmic pattern* and *meter*. A rhythmic pattern can be most simply defined as a series of temporal intervals. Figure 1.1 (top) depicts a rhythmic pattern in musical notation and in graphical form, made up of a distinctive series of short and long temporal intervals having 250 ms and 500 ms durations. Meter, in contrast, is the abstract temporal structure of music, composed of multiple nested periodic structures, the most salient of which is experienced as the "beat," the point in time where most people tap their fingers or feet to the music. Unlike rhythmic pattern structure, which corresponds to the specific pattern of temporal intervals in a sequence, metrical structure must be inferred from periodic regularities in the musical surface (Clarke, 1999; Palmer & Krumhansl, 1990). Listeners tend to infer from music a primary, periodic cycle marked by equally spaced (i.e., "isochronous") beats, plus one or two faster or slower isochronous levels that subdivide or multiply the primary cycle. Two common types of metrical subdivisions are *duple*, in which the primary cycle is subdivided into four or two beats, and *triple*, in which it is subdivided into three beats. This hierarchical structure creates alternating patterns of perceived strong and weak beats in music, as in the triple meter waltz pattern of "**one** two three, **one** two three…" illustrated in Figure 1.1. Note that during first cycle of the meter the rhythmic pattern is composed of three short intervals, while during the second cycle it is composed of a long interval followed by a short interval; this illustrates that the rhythmic pattern can and often does vary independent of the meter. Metrical structure enables individuals in a group to synchronize their movements in dancing, marching, tapping, clapping, and singing, allowing for precise anticipation of when movements should occur. Metrical structure is also thought to guide attention, enhancing anticipation for what is likely to occur and when it will occur, and aiding in segmentation of musical sequences into melodic and rhythmic groups (Jones & Boltz, 1989; Large & Jones, 1999; Palmer &

Pfordresher, 2003). Because synchronized coordination of movement to music has been observed in all known cultures, meter is thought to constitute a fundamental and universal aspect of musical perception and behavior (Brown, 2003).

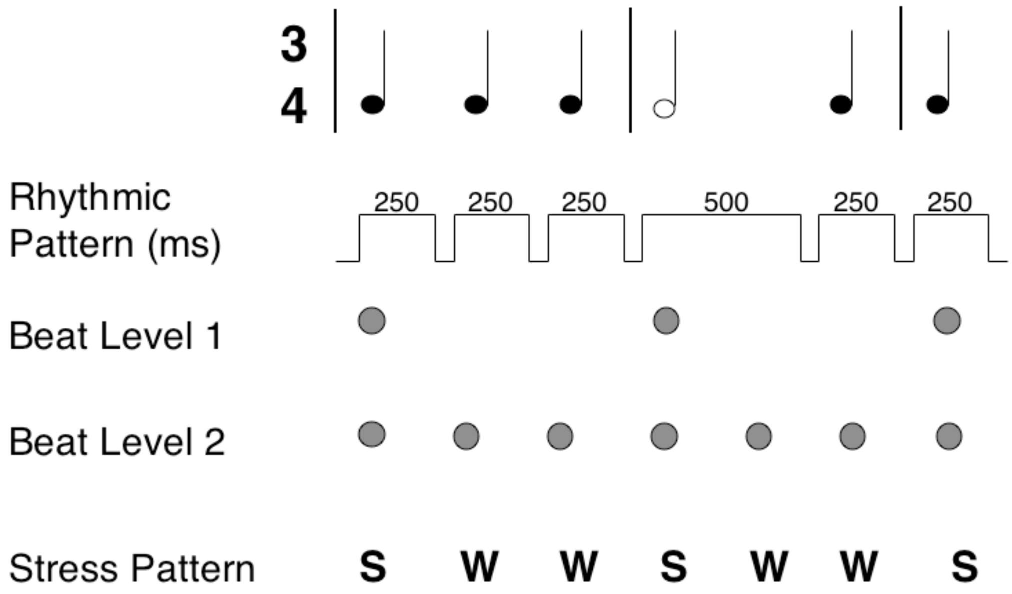


Figure 1.1. A metrical hierarchy typical of a waltz is depicted in musical notation and graphical form. For the triple meter time signature 3/4, the denominator specifies the base unit (a quarter-note) and the numerator specifies the number of units per cycle. Alternating strong (S) and weak (W) beats arise from the metrical hierarchy, which consists of two isochronous beat levels.

It is difficult to predict exactly how a musical sequence will give rise to a subjective pattern of periodic strong and weak beats, but distributional regularities are likely to serve as essential cues. In Western music, composers use time signature notation to specify the intended meter. For example, the time signature 3/4, which corresponds to the meter depicted in Figure 1.1, signifies that the primary metrical cycle is subdivided into three beats, the first of which has the greatest metrical strength. For a given meter, note onsets, or events, tend to occur more frequently at strong (downbeat) rather than weak (upbeat) positions in the primary metrical cycle. The frequency of event occurrence has been shown to predict the meter of musical excerpts from four styles of Western classical music (Palmer & Krumhansl, 1990) and children's nursery tunes (Palmer & Pfordresher, 2003). Thus, musical events, such as the sounding of a piano or the beating of a drum, are much more likely to occur every second or fourth beat in duple meters and every third beat in triple meters. A second type of distributional cue to meter is *accent.* Accents arise when events are relatively salient, due to being longer, louder, higher in pitch, positioned at points of change in a melody, or at melodic and/or rhythmic group boundaries. For the remainder of the paper, our use of the term *accent* will denote such *phenomenal* accents, to be distinguished from *metrical* accents, which arise from the pattern of strong and weak beats that are perceived after a metrical representation has been activated (Lerdahl & Jackendoff, 1983). Analyses of various styles of western music reveal that accents tend to occur frequently at metrically strong positions in a given time signature (Huron & Royal, 1996; Longuet-Higgins & Lee, 1982). The frequency of events and accents is thus a reliable distributional cue to meter.

The widespread occurrence of synchronized dancing, tapping, and other types of rhythmic behavior attests to the relative ease with which adults can grasp the intended metrical structure of most music. Several studies have shown that adults accurately synchronize their tapping with strong metrical positions in music, exhibiting a high level of inter-subject agreement (Drake, Penel, & Bigand, 2000; Snyder & Krumhansl, 2001; Toiviainen & Snyder, 2003). Frequency of accents at potential downbeat positions reliably predicts tapping position (Snyder & Krumhansl, 2001) and the perceived meter (Hannon, Snyder, Eerola & Krumhansl, 2004). Adults thus appear to use the distribution of accents and events over time to infer metrical structure.

Unlike adults, infants are not capable of producing precisely timed rhythmic behaviors such as synchronized tapping. However, their rhythmic pattern perception parallels that of adults in important respects. Like adults, 2- and 5-month-old infants can use the relative size and order of intervals to discriminate rhythmic patterns, such as 600-200-400 ms vs. 200-600-400 ms (Demany, McKenzie, & Vurpillot, 1977; Chang & Trehub, 1977). Seven- and 9-month-old infants can categorize stimuli on the basis of identical rhythmic pattern structure, despite concurrent transformations of frequency and tempo (Trehub & Thorpe, 1989, but see Pickens & Bahrick, 1997). Thus adults and infants both tend to perceive rhythmic patterns in terms of global and relational information and not in terms of the component temporal intervals that compose a pattern (Trehub, Trainor, & Unyk, 1993).

Infants' abilities for discriminating and categorizing stimuli on the basis of rhythmic pattern structure can be considered a prerequisite for perceiving meter, because rhythmic patterns are thought to guide meter induction in important ways. For example, temporal accents can arise from the position of an event in relation to the rhythmic grouping of a pattern, or the way in which events are clustered in time (Lerdahl & Jackendoff, 1983). An event tends to sound accented if it is surrounded by silence (isolated), or if it falls at a boundary of a group of events surrounded by silence (Povel & Essens, 1985). Thus, if infants are sensitive to the rhythmic grouping of events, they may also hear subjective temporal accents that might enable them to infer meter. Nevertheless, only indirect evidence suggests that infants perceive meter.

Metrical information is salient in the environment of young infants, who are frequently rocked and bounced to music (Papousek, 1996). When mothers sing to their infants, they tend to emphasize metrically stressed syllables through small increases in duration and loudness (Trainor, Clark, Huntley, & Adams, 1997). Such changes are noticed by 10-month-old infants, who distinguished otherwise identical musical segments performed in different metrical contexts (Palmer, Jungers, & Jusczyk, 2001). As early as 2 months, infants can detect changes to the speed of an isochronous pattern, which might reflect a rudimentary form of meter perception (Baruch & Drake, 1997). Nine-month-old infants detect temporal deviations to rhythmic patterns better when those patterns induce a strong metrical framework (Bergeson, 2002). These findings indirectly suggest that infants perceive meter. A primary goal of the present work was to directly tackle this question by measuring whether infants could categorize unique rhythms on the basis of a common underlying meter. Such a result would provide the strongest evidence to date that infants perceive metrical structure in musical patterns.

A secondary goal was to investigate whether meter, if perceived by infants, might serve to facilitate other aspects of music learning. Studies of speech perception have shown that distributional cues can interact in important ways, with certain types of cues providing a foundation for learning other types of information. For example, stress is a reliable distributional cue to word boundaries in English because most words begin with stressed syllables. At 7 to 9 months of age, infants exploit this regularity to segment words with initial stress from fluent speech (Curtin et al., in press; Jusczyk, Houston, & Newsome, 1999). Unlike 7-month-olds, 9- to 11-month-old infants can correctly segment words having the atypical weak-strong pattern (Jusczyk et al., 1999; cf. Thiessen & Saffran, 2003). The superior performance of older infants presumably results from the acquisition of additional knowledge about word boundaries that overrides stress, such as *phonotactic* regularities, which are the sequences of speech sounds that are typical within versus between words (Jusczyk et al., 1999; Myers, Jusczyk, Kemler-Nelson, Charles-Luce, Woodward, & Hirsh-Pasek, 1996). Although various cues to word boundaries are probably learned simultaneously, an initial tendency to associate stressed versus unstressed syllables with different speech sounds may highlight word boundaries, facilitating acquisition of knowledge about other segmentation cues in language (Christiansen et al., 1998; Curtin et al., in press; Jusczyk, 1999). A similar type of bootstrapping process may exist in musical structure learning (Jones, 1990). If infants can use the distribution of event and accent occurrence to infer the meter, the meter may in turn serve as a framework for learning about other aspects of pitch structure in music by highlighting particular events in a sequence.

To summarize, little is known about whether infants perceive the underlying temporal structure of music, the meter. Distributional information, such as frequency of event and accent occurrence, reliably predicts the meter perceived by adults. Infants' general sensitivity to distributional cues in auditory patterns, as well as their adult-like perception of rhythmic patterns, suggests that they might be able to use event and accent occurrence to infer meter. In the first two experiments we examine whether infants can infer meter from simple rhythmic patterns. If they can, they should be able to differentiate new rhythmic patterns that conform to the inferred meter from those that do not. In the third experiment, we examine whether infants can learn to associate musical pitches with strong or weak positions in the meter.

1.2. Experiment 1

The first experiment assessed whether infants could use distributional cues to infer the underlying meter of simple rhythmic patterns. Previous studies have shown that infants respond similarly to unique instances of a stimulus if those instances have the same structure or fall into the same category (Trehub & Thorpe, 1989). If infants perceive metrical structure, they should respond similarly to a set of unique rhythmic patterns that induce the same meter. In Experiment 1, 7-month-old infants were habituated to an audio-visual display that presented three different rhythms having the same underlying meter. Following habituation, infants were presented with two novel rhythms that implied a novel or a familiar meter. If infants perceive meter, they should exhibit differential looking for rhythmic patterns that imply a novel vs. familiar meter.

1.2.1. Method

1.2.1.1. Participants

Twenty-four 7-month-old infants (*M* age = 217.7 days, *SD* = 4.4) participated, 11 girls and 13 boys. Three additional infants were tested but not included in the sample due to fussing (*N* = 3), or experimenter error (*N* = 1). All participants were healthy, full-term infants with no complications during delivery and no reported health or hearing problems.

1.2.1.2. Stimuli

Eight rhythms were composed of 24 temporal units that were 250 ms in duration, 9 of which were *silent* units and 15 of which were *event* units. Event units consisted of a 100 ms tone at the pitch level of C5 (523 Hz), followed by 150 ms of silence. The brief duration of tones gave them a staccato quality. Tones were generated using Quicktime's *Ocarina* timbre, which approximates a sine tone. Silent units were 250 ms in duration. Two consecutive event units resulted in a 250 ms inter-onset interval, while two consecutive silent units resulted in a cumulative silence of 500 ms. The longest possible silent duration was 500 ms, and this occurred at least once in each rhythm. Figure 1.2 depicts the temporal properties of a brief segment of one rhythm in iconic form, with event units designated by "x" and silent units designated by "o".

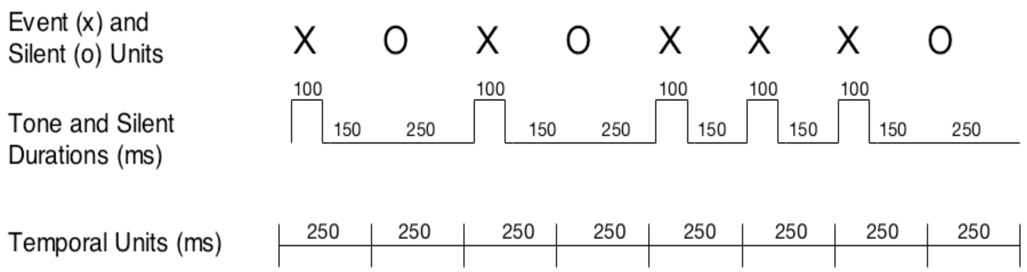


Figure 1.2. An iconic depiction of a rhythm segment. Temporal units of 250 ms were either silent units (250 ms, depicted as "o"), or event units (100 ms tone followed by 150 ms silence, depicted as "x").

Rhythms were created to imply either a duple meter or a triple meter. In triple meter, events and accents occurred more frequently on the first of every three units (and less frequently at all other positions), and in duple meter they occurred more frequently on the first of every four or two units. To minimize differences between rhythms, we avoided physically altering the amplitude or pitch level of accented versus unaccented events, instead focusing on the subjective accents that arise from the positioning of events relative to rhythmic groups. We hypothesized, based on previous adult research, that events would sound accented if they were relatively isolated, the second of a two-event group, or the first or last event in a group larger than three (Povel & Essens, 1985). We thus manipulated the frequency distribution of accents and events by assigning an event or silence to each temporal unit quasi-randomly, with the following constraints adapted from Povel and Essens (1985): (a) No silences occurred at strong metrical positions, (b) events in strong positions could not be both preceded and followed by other events, (c) events in weak positions could not be followed by silence. The first constraint ensured that events occurred more frequently at strong metrical positions, while the other constraints ensured that accents occurred more frequently at strong than at weak metrical positions by keeping strong events isolated or at group boundaries (b) and preventing weak events from being relatively isolated or occurring at the end of a group boundary (c). For triple meter rhythms, every third unit was designated as metrically strong. Because theoretical descriptions of duple meter distinguish between the primary downbeat, which would occur every four units, and the secondary downbeat, which would occur every two units, we used all four constraints for primary downbeat units (every four), and constraint (c) for secondary downbeat units (every two). This resulted in a slightly lower frequency of accents and events at secondary versus primary metrical positions.

Duple Rhythms

D1: | **x** o **x** o | **x** x **x** o | **x** o o x | **x** o **x** o | **x** x x x | **x** o o x |

D2: | **x** o o x | **x** o o x | **x** o **x** x | **x** o **x** o | **x** x x x | **x** o **x** o |

D3: | **x** x **x** o | **x** x x x | **x** o o x | **x** o **x** o | **x** o o x | **x** o **x** o |

D4: | **x** o o x | **x** o **x** o | **x** x x x | **x** o **x** x | **x** o o x | **x** o **x** o |

Triple Rhythms

T1: | **x** x x | **x** o x | **x** o x | **x** o o | **x** o x | **x** o o | **x** x x | **x** o o |

T2: | **x** o o | **x** x x | **x** o x | **x** o o | **x** o o | **x** x x | **x** o x | **x** o x |

T3: | **x** o x | **x** o x | **x** o o | **x** x x | **x** o o | **x** o x | **x** o o | **x** x x |

T4: | **x** x x | **x** o x | **x** o o | **x** x x | **x** o x | **x** o o | **x** o x | **x** o o |

Figure 1.3. Rhythmic patterns used in Experiment 1. x = event unit, o = silent unit. Accents are indicated in bold.

Using these constraints, we generated four unique rhythms in each meter (Figure 1.3). Rhythms were differentiated by the distribution of events and accents (in bold) at strong vs. weak positions in triple or duple meter. Figure 1.4 displays the average proportion of events and accents that occurred for every two, three, or four units of the duple and triple meter stimuli. As intended, duple meter stimuli had a higher proportion of events and accents occurring every two or four units, while triple meter stimuli had the highest proportion of events and accents occurring every three units.

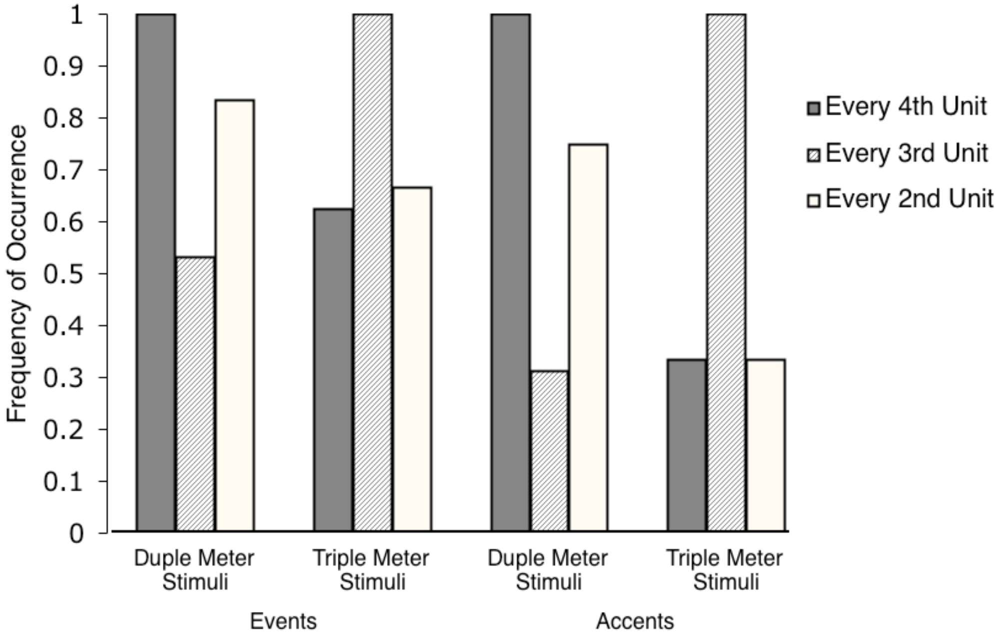


Figure 1.4. The average proportion of event and accent occurrence is shown for downbeats in duple meter (every 4th unit), downbeats in triple meter (every 3rd unit) and secondary downbeats in duple meter (every 2nd event). The proportion was calculated by dividing the number of times an accent or event occurred by the total number of potential downbeat units in each rhythm.

Even though previous findings suggested that adults use frequency of event and accent occurrence to infer the metrical structure (Hannon et al., 2004; Povel & Essens, 1985; Snyder & Krumhansl, 2001), we collected pilot data to ensure that our stimuli sounded metrical to adults. Three expert musicians[1] rated each rhythm's degree of fit to both triple and duple meter. We subtracted the average ratings for triple fit from those for duple fit to obtain a judgment scale ranging from strong duple fit (positive) to strong triple fit (negative). As predicted, ratings were significantly higher for the four rhythms we had designated duple than for the four rhythms we had designated triple, $p < .01$. In other words, duple rhythms sounded much more "duple" than triple rhythms and vice versa. Because accented events were physically indistinguishable from non-accented events, we also wanted to verify that intended accents were perceived as such by asking musicians to mark which events sounded accented. The average match between perceived and intended accents was 90%, indicating that most accents were reliably perceived. We also assessed whether non-musically trained adults could differentiate the stimuli on the basis of implied meter. In a paired comparison task, adults indicated which of two novel comparison rhythms had the same underlying beat as two standard rhythms with 75% accuracy, which was significantly above chance, $p < .001$. Although performance was only moderately accurate, we suspect the task was somewhat difficult because of adults' known tendency to interpret rhythms in duple meter, perhaps due to the greater prevalence of duple meters in Western music (Hannon et al., 2004).

Each 6 s rhythmic pattern was cycled ten times to create a maximum trial duration of 60 s. All rhythms were combined with a video display and converted into QuickTime movies. The video display consisted of an unmoving black and white

---

[1] Three musicians had an average of 25 years music lessons and advanced music degree certification and/or training from Moscow State Conservatory College of Music, Royal Conservatory of Toronto, and Eastman School of Music.

checkerboard that filled the screen and remained present throughout the duration of each trial.

1.2.1.3. Apparatus

A Macintosh G4 computer and a 76 cm monitor equipped with a speaker were used to present audio-visual stimuli and collect looking time data. A camera placed on top of the monitor recorded the infant and transmitted the image to a second monitor behind a large barrier. The experimenter viewed the infant on the second monitor and entered judgments of infant gaze with a key press on a computer keyboard. Because the monitor speakers presented the rhythms, infant head turns towards the monitor also reflected turns toward the sound source. The computer presented the audio-visual stimuli (i.e., the checkerboard and rhythmic patterns), recorded looking time judgments, and calculated habituation criteria for each infant. The experimenter and parent wore headsets playing music to mask the stimulus sounds.

1.2.1.4. Procedure

Infants were tested individually and sat on their parents' laps in a darkened room, at a distance of 90 cm and turned at a 45° angle to the left of the monitor. This angle required a slight head turn towards the monitor. All trials were initiated as soon as the infant fixated on the monitor, and terminated when the infant looked away for more than 2 s. Between trials, the computer presented a looming target accompanied by a rapidly pulsed siren to orient the infants' attention towards the monitor.

During the habituation phase, each infant was presented with a series of three rhythms all in the same meter. The order of rhythm presentation on habituation trials was quasi-random, with the restriction that all three rhythms had to be presented before a given rhythm could be repeated. Each infant was assigned to one of two habituation conditions, in which the meter of the habituation rhythms was duple or triple. Infants were habituated to the sequence of three rhythms until habituation of looking occurred or 12 trials had elapsed. The habituation criterion was defined as an average fixation decrement of 50% over four trials relative to the average fixation of the previous four trials.

Immediately following habituation trials, infants were presented with six test trials, consisting of three alternating presentations of two novel rhythms, one from a novel meter and one from a familiar meter. Order of post-habituation test trials was counter-balanced in each condition, so that half of the infants were presented with a novel meter rhythm first, and half were presented with a familiar meter rhythm first. The rhythms used during habituation vs. test phase were alternated throughout the experiment, so that all rhythms occurred during both the habituation and test phases of the experiment.

1.2.2. Results and Discussion

Infants oriented longer towards the post-habituation test rhythm that implied a novel meter than towards the test rhythm that implied a familiar meter (see Figure 1.5). Looking time data were positively skewed in some cells, so all data were log-transformed prior to analyses (data shown in Figure 1.5 are raw scores). A three-way, mixed design ANOVA, with test condition (novel vs. familiar meter, within subjects), habituation condition (triple vs. duple meter, between subjects), and test order (novel meter first vs. familiar meter first, between subjects), revealed a significant main effect of test condition, $F(1, 20) = 7.05$, $p < .05$. There were no other significant main effects or interactions with habituation condition or test order.

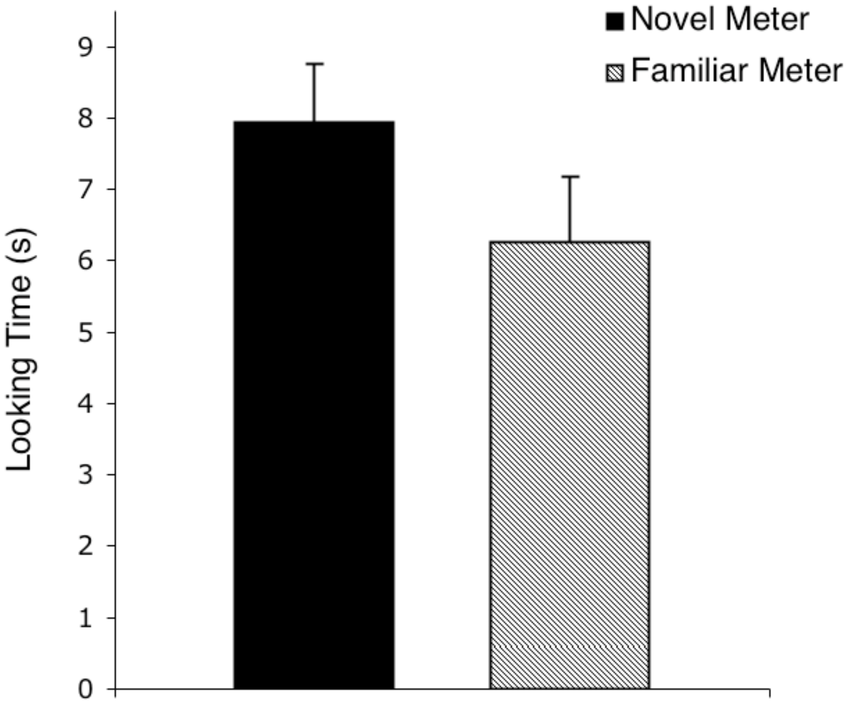


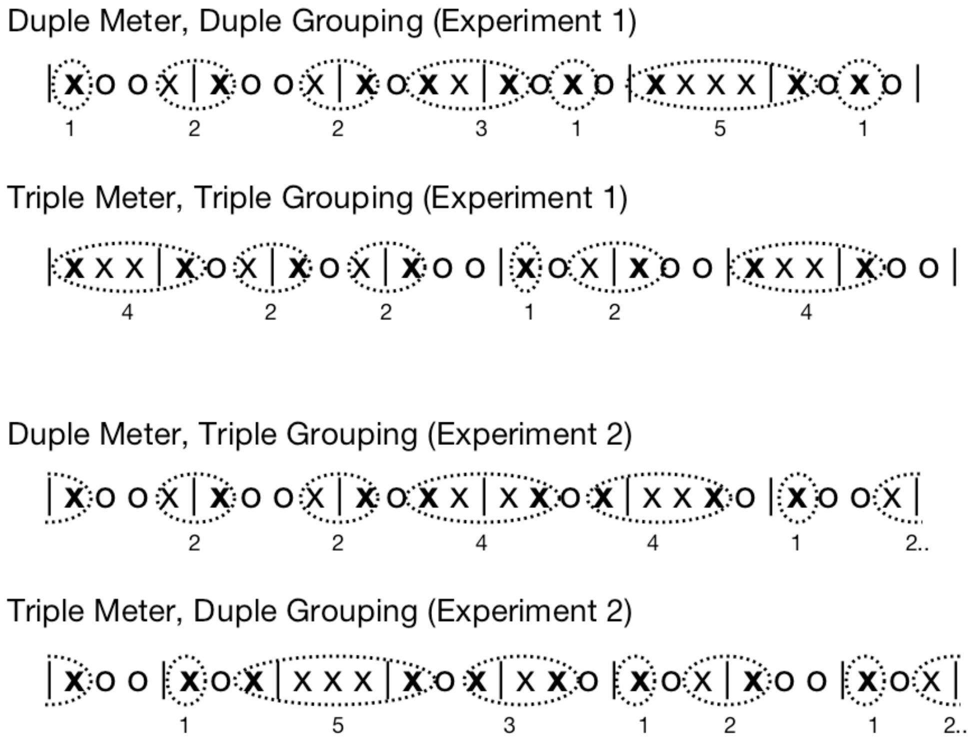
Figure 1.5. Mean looking times after habituation (in seconds). Error bars indicate standard errors. Infants looked longer during presentation of a novel meter rhythm.

This result suggests that infants inferred the underlying meter from the three rhythms presented in the habituation phase, which resulted in a novelty preference for the test rhythm that induced a novel meter. Because all rhythms were identical in length and in the number of events and silences, it is likely that the distribution of event and accent occurrence was the primary differentiating feature of triple versus duple meter rhythms. We assume that infants inferred periodic accents from the positioning of events relative to rhythmic groups, but it is possible that infants simply noticed and remembered the grouping structure (the number and size of groups).

Because group size was also a differentiating feature of these rhythms, we conducted a second experiment to separately assess the effects of grouping structure versus implied meter on infant behavior.

1.3. Experiment 2

The triple and duple meter stimuli used in Experiment 1 were differentiated by the distribution of events and accents as well as by the number of events per group. We define a group as any series of events bordered on both sides by silence. An unintended outcome of the constraints used to generate stimuli in Experiment 1 was that some group sizes occurred exclusively in one meter and not in the other. In Figure 1.6, all groups in four rhythms have been circled and labeled according to size, i.e., the number of events within each group. Notice that for Experiment 1 (Figure 1.6 top and Figure 1.3), duple rhythms contained groups of 1, 2, 3, and 5, while triple rhythms contained groups of 1, 2, and 4. Groups of three events, for example, occurred in duple but not triple meter rhythms. This is because groups of three events were inherently problematic for generating a metrically appropriate distribution of events and accents. To illustrate, a group of three events starting at a downbeat position in triple meter would leave the subsequent downbeat silent, which violates the constraint that events should always occur at strong metrical positions.


Duple Meter, Duple Grouping (Experiment 1)

Triple Meter, Triple Grouping (Experiment 1)

Duple Meter, Triple Grouping (Experiment 2)

Triple Meter, Duple Grouping (Experiment 2)

Figure 1.6. Rhythmic patterns used in Experiment 2. x = tone, o = silence. Each group is circled with a dashed line, and group sizes are labeled beneath. Accents are indicated in bold.

Previous studies have shown that 12-month-old infants can discriminate rhythmic patterns on the basis of group size differences (Morrongiello, 1984), so it is possible that infants categorized rhythms in Experiment 1 on the basis of group size. This possibility would be consistent with the hypothesis that infants process rhythmic patterns according to serial structure, and that they have not yet developed the ability to infer the complex hierarchical aspects of metrical structure (Drake, 1998). To disentangle the effects of grouping structure and meter, we replicated Experiment 1 using test rhythms that pitted grouping structure against implied meter.

1.3.1. Method

1.3.1.1. Participants

Twenty-four 7-month-old infants (*M* age = 220.4 days, *SD* = 15.9) participated, 14 girls and 10 boys. Two additional infants were observed but not included in the sample due to fussing. All participants were healthy, full-term infants with no complications during delivery and no reported health or hearing problems.

1.3.1.2. Stimuli

The same rhythms from Experiment 1 were used during the habituation phase. Two additional rhythms were created for use during the test phase. As in Experiment 1, test rhythms were composed of 24 temporal units, made up of 9 silences and 15 tones. The pitch, timbre, durations, and inter-onset intervals of the rhythms were identical to those used in Experiment 1.

In Experiment 1, all rhythms in a given meter were identical in grouping structure. As shown in Figure 1.6 and Figure 1.3, duple meter rhythms always contained three groups of one, two groups of two, one group of three, and one group of five, while triple meter rhythms contained one group of one, three groups of two, and two groups of four. To create test rhythms for Experiment 2, the grouping structure of each meter was rearranged according to the metrical constraints of the contrasting meter, to the maximum extent possible. Therefore, the triple meter test rhythm contained group sizes identical to that of duple meter habituation rhythms (triple meter, duple grouping), while the duple meter test rhythm had grouping structure identical to that of triple meter habituation rhythms (duple meter, triple grouping). Both test rhythms are presented in Figure 6 (bottom).

As explained above, certain group sizes created metrical ambiguity, such as groups of three in triple meter. Figure 1.7 presents the proportion of events and accents at potential downbeat positions in each triple and duple meter test rhythm.

Compared to the stimuli from Experiment 1, these triple and duple meter rhythms are not as strongly differentiated by frequency of event or accent occurrence. It is important to point out, however, that rhythms need not provide perfect information to imply one meter or the other. Because meter is inferred from probabilistic information, some ambiguity should be tolerable if, overall, a rhythm is consistent with a particular meter. A real world example is *syncopation*, a phenomenon in which events and accents occur “off” the beat, creating tension but not disrupting the perception of meter altogether, especially once listeners have inferred a metrical framework (Clarke, 1999).

Thus each test rhythm presented features that were both novel and familiar relative to habituation rhythms. One stimulus presented a novel grouping structure but a familiar implied meter, while the other presented a familiar grouping structure but a novel implied meter.

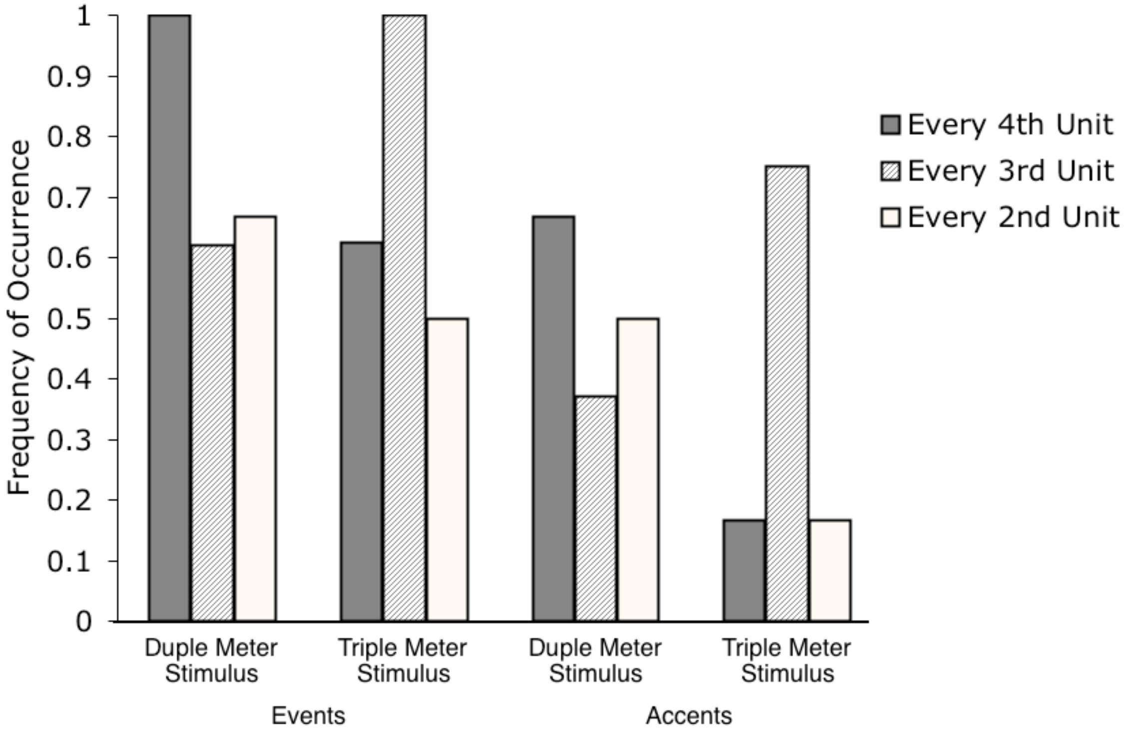


Figure 1.7. The average proportion of event and accent occurrence for rhythms used in the test phase of Experiment 2.

1.3.1.3. Apparatus and Procedure

The apparatus and procedure were identical to Experiment 1 with the following exceptions. During the habituation phase, infants were presented with a series of three of the rhythms used in Experiment 1, all in the same meter (see Figure 1.3). Immediately following the habituation phase, infants were presented with two novel test rhythms, one with familiar grouping/novel meter and the other with novel grouping/familiar meter. As in Experiment 1, the meter of habituation rhythms and the order of test trials was counter-balanced.

1.3.2. Results and Discussion

Infants oriented longer to presentation of post-habituation test rhythms that implied a novel meter and familiar grouping structure than to rhythms that implied a familiar meter and novel grouping structure (raw looking time data are presented in Figure 1.8). A three-way ANOVA, with test condition (novel meter/familiar grouping vs. familiar meter/novel grouping, within subjects), habituation condition (triple vs. duple meter, between subjects), and test order (novel meter first vs. familiar meter first, between subjects), revealed a significant main effect of test condition, $F(1, 20) = 4.41$, $p < .05$. There were no other significant main effects or interactions with habituation condition or test order.

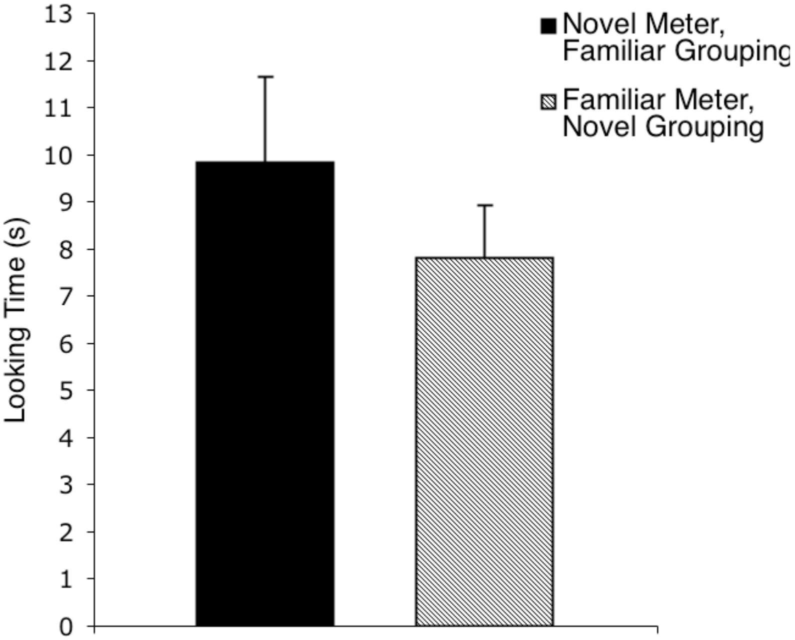


Figure 1.8. Mean looking times after habituation (in seconds). Error bars indicate standard errors. Infants looked longer during presentation of the rhythm that implied a novel meter but a familiar grouping structure.

This result indicates that infants categorized rhythms on the basis of implied meter and not on the basis of grouping structure. If infants had responded to novel grouping structure in Experiment 1, they should have also shown a preference for novel grouping structure in Experiment 2. However, they showed a preference for the familiar grouping structure and the novel metrical structure. Moreover, because we used infant-controlled habituation, we can be confident that the results obtained in both Experiments 1 and 2 reflect a novelty preference (Horowitz, Paden, Bhana, & Self, 1972). In the Experiment 2, test stimuli each presented a different type of novel structure, allowing us to determine which aspect of habituation rhythms dominated infant perception. We can thus conclude that infants we observed in Experiments 1 and 2 inferred the underlying meter from the distribution of events and accents in the set of habituation rhythms and not from common grouping structure. Combined, the first two experiments provide the strongest evidence to date that infants can infer metrical structure from rhythmic patterns.

Our findings suggest that infants can perceive meter even though they are not yet capable of producing precisely timed movements in synchrony with music. These results are surprising because they may document a precocious grasp of hierarchical temporal structure relative to hierarchical pitch structure in music, which is not grasped until late childhood (Krumhansl & Keil, 1982; Trainor & Trehub, 1994; Wilkes, Wales, & Pattison, 1997). This ability raises questions about the potential functions of metrical structure. Periodic temporal structures could function to draw attention to particular events or relationships between events, thus forming a basis for learning about pitch relations in music (Jones, 1990). It is possible that early abilities for perceiving meter might guide infant attention, enabling infants to interpret and organize incoming musical information. Just as stress facilitates learning of phonotactic cues in speech segmentation, meter might facilitate infants' learning of the complex and hierarchical pitch structures in music.

Tonality, sometimes called the "syntax" of music, refers to the hierarchical system of pitch relations in Western music, specifying concepts of *scale*, *chord*, and *key* at successive hierarchical levels (Cuddy & Badertscher, 1987; Patel, 2003). A large body of evidence suggests that adults possess tacit knowledge of tonality, which allows them to perceive the relative prominence of pitches and to detect “sour” notes in conventional musical contexts (Krumhansl, 1990). Individuals likely acquire this knowledge some time after infancy, as shown by striking differences between infant and adult performance on tasks measuring sensitivity to scale structure (Lynch, Eilers, Oller, & Urbano, 1990), chords (Trehub, Cohen, Thorpe, & Morrongiello, 1986), and key (Trainor & Trehub, 1992). Very little is known about how individuals acquire tonal knowledge, but some evidence suggests that frequency of occurrence and final note position can determine adults’ inferences of which pitches are prominent in an unfamiliar pitch system (Creel & Newport, 2002).

Music theorists have asserted that structurally important events occur more frequently at metrically strong positions (Meyer, 1973). This hypothesis has been supported by some empirical findings. An analysis of jazz improvisations showed that musicians tended to play structurally important notes at metrically strong locations (Jarvinen & Toiviainen, 2000). An analysis of errors by adult and child pianists revealed that notes at strong metrical positions were more likely to be replaced with notes from other strong metrical positions than with notes from weak positions, suggesting that pianists conceptualize pitches according to their position in the meter (Palmer & Pfordresher, 2003). These findings lend support to the proposal that metrical structure, by emphasizing some pitches over others, might serve as a cue for learning complex aspects of pitch structure such as tonality. If meter facilitates infants’ learning of hierarchical pitch relations in music, infants should be able to learn an association between strong vs. weak metrical positions and the pitches that tend to occur at those positions.

1.4. Experiment 3

A final experiment aimed to investigate one potential function of metrical structure. The design from Experiments 1 and 2 was adapted for use with melodies. We created two *tone distributions* (which we labeled A and B) that differed in the frequency with which certain pitches occurred at metrically strong vs. weak positions. For melodies having tone distribution A, half of the pitches occurred more frequently at metrically strong positions, and the other half of pitches occurred more frequently at metrically weak positions. For melodies having tone distribution B, the opposite set of pitches occurred at strong vs. weak metrical positions. During the habituation phase, infants were presented with a series of melodies having the same tone distribution. During the test phase, they heard two novel melodies having a novel or a familiar tone distribution. In a control condition, infants were presented with non-rhythmic (isochronous) versions of the same melodies. If infants learned to associate certain pitches with metrically strong and/or weak positions, we expected infants in the experimental but not the control group would show a novelty preference for melodies having a novel tone distribution.

1.4.1. Method

1.4.1.1. Participants

Forty-eight 7-month-old infants (*M* age = 213.7 days, *SD* = 12.8) participated, 24 girls and 24 boys. Five additional infants were observed but not included in the sample due to fussing or interruption during testing. All participants were healthy, full-term infants with no complications during delivery and no reported health or hearing problems.

1.4.1.2. Stimuli

Sixteen rhythmic melodies were created for the experimental condition, and sixteen isochronous melodies were created for the control condition.

1.4.1.2.1. Experimental Stimuli

For the experimental condition, triple meter rhythmic patterns from Experiment 1 were adapted for use as melodic stimuli. Instead of each tone in the rhythm consisting of a fixed pitch, however, each event was assigned one of six pitch values: C4, D4, E4, F#4, G#4, and A#4. These particular pitches were chosen because they are members of a *whole-tone* scale. Whole-tone scales have pitches that are all equally separated by two semitones. We used the whole-tone scale because it is relatively uncommon in Western music and was unlikely to have been encountered by infants. In addition, whole-tone scales lack the perceptually prominent perfect fifth interval, which if present could potentially bias learning in some conditions and not others.

Tone distributions A and B can be conceptualized as artificial tonalities, differentiated by the frequency with which certain pitches occur at strong versus weak metrical positions. To create tone distribution A, the six pitches were randomly assigned membership to either a strong or a weak group. Pitches assigned to the strong group occurred at strong metrical positions 90% of the time and at weak positions 10% of the time. The reverse was true of pitches assigned to the weak group, which occurred at strong metrical positions only 10% of the time and at weak positions 90% of the time. The pitch of a given event in a rhythm was determined pseudo-randomly. If an event occurred at a strong metrical position in the rhythm it had a 90% chance of being assigned one of the three pitches in the strong group, and events at weak positions had a 90% chance of being assigned a pitch in the weak group. To create tone distribution B, each pitch's assignment was reversed; pitches in the strong group were switched to the weak group and vice versa. Thus, eight unique melodies were created, four for each tone distribution. Analogous to Experiment 1, three melodies within the same category (tone distribution A, for example) could be presented during the habituation phase, while the fourth melody from that familiar category (distribution A) could be paired with a melody from a novel category (distribution B) during the test phase.

To control for the possibility that melodies would be differentiated by the overall frequency of occurrence for individual pitches or sets of pitches regardless of the meter, each melody had a frequency-matched melodic counterpart with an opposite tone distribution but an identical number of occurrences for each pitch. To illustrate, if pitch D occurred six times in a melody from tone distribution A, its melodic counterpart from tone distribution B also contained six instances of pitch D. The two melodies differed only in the placement of that pitch with respect to the meter, with D occurring most often on strong beats in distribution A but on weak beats in distribution B. Figure 1.9 shows the proportion of time each pitch occurred at strong vs. weak metrical positions in two frequency-matched melodies, one with distribution A and one with distribution B. When melodies were presented during the test phase, each melody could thus be presented with its frequency-matched counterpart differing only in its tone distribution.

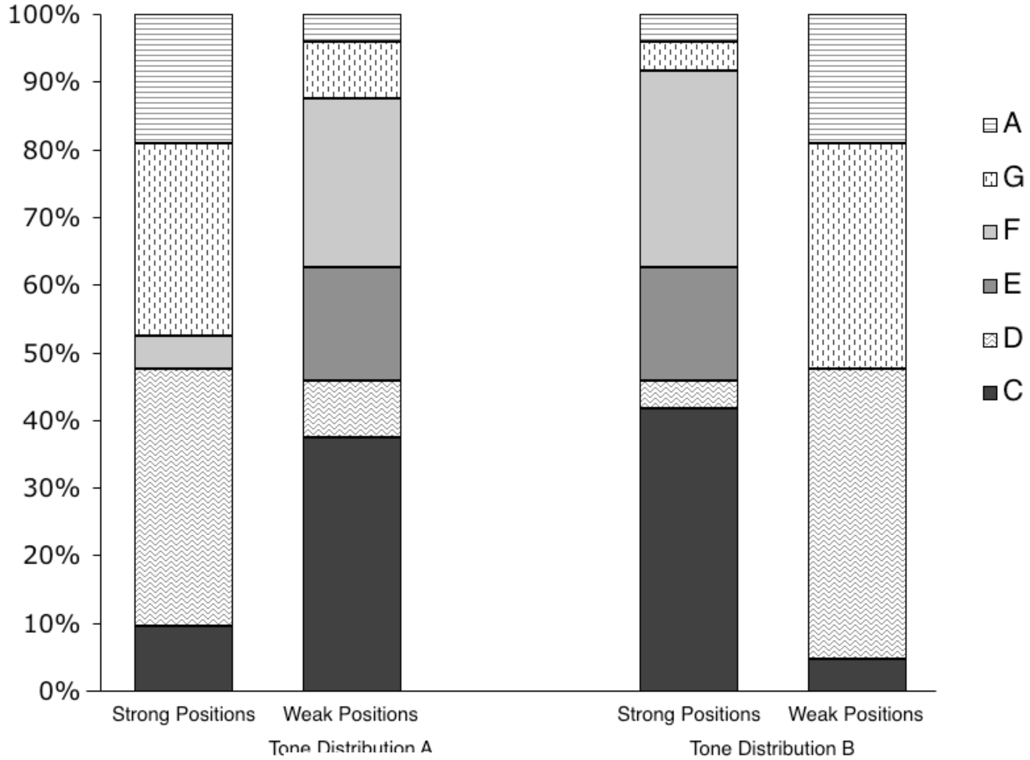


Figure 1.9. One pair of frequency-matched melodies from two different tone distributions. Proportions indicate the percentage of time each pitch occurred at strong vs. weak metrical positions in the distributions A and B. In this example, pitches A, G, and D occurred more frequently at strong than at weak metrical positions for tone set A, while they occurred more frequently at weak than at strong positions for tone set B.

Assignment of individual pitches to strong vs. weak groups was random, but some configurations might have unintentionally resulted in distinctive melodic features, such as a predominance of rising or falling pitch contours for one tone distribution but not the other. To minimize the possibility of creating such artifacts, we generated each set of tone distributions twice using different *group assignments* as an additional control. For group assignment, pitches D, G#, and A# composed a group (strong in tone distribution A) while pitches C, E, and F# composed the other group (strong in tone distribution B). For the other group assignment, pitches C, E, and G# composed one group (strong in A) while pitches D, F#, and A# composed the other (strong in B).

We assigned pitches to the first 45 events of triple meter rhythms from Experiment 1 to create a total of 16 unique melodies (four melodies with tone distribution A plus four melodies with distribution B, generated twice according to two different group assignments). Each 18 s melody was cycled four times to create a maximum trial duration of 72 seconds. The timbre, durations, and inter-onset intervals of rhythms were identical to those used in Experiment 1. Although the rhythmic patterns from Experiment 1 contained regular temporal grouping accents, unintended melodic accents in the present stimuli could influence or confuse meter induction (Hannon et al., 2004). To make the meter unambiguous, all trials were preceded by a brief four-cycle drum lead in, which established the primary metrical beat.

1.4.1.2.1. Isochronous Control Stimuli

We wanted to rule out the possibility that infants might differentiate test melodies on the basis of sequential regularities in pitch structure, unrelated to meter. We therefore created a set of isochronous melodies for use in a control condition. For each unique melody used in the experimental condition, an isochronous version was created that consisted of an identical sequence of pitches but no rhythmic variation (i.e., no silent units). Control stimuli had 45 isochronous events having a 250 ms inter-onset interval, cycled six times to create a maximum trial duration of 67 s.

1.4.1.3. Apparatus and Procedure

The apparatus and procedure were identical to Experiments 1 and 2 with the following exceptions. The habituation phase consisted of a rotating series of three melodies, all with the same tone distribution. Immediately following the habituation phase, infants were presented with two novel melodies, one with a novel tone distribution and the other with a familiar tone distribution. In the test phase, the melody with the familiar tone distribution was always presented with its frequency-matched counterpart, so individual pitches occurred with equal frequency in both test melodies. Half of the infants were assigned to the experimental condition and half were assigned to the isochronous control condition. Group assignment, habituation condition (tone distribution A or B), and test trial order were counter-balanced.

1.4.2. Results and Discussion

A four-way mixed-design ANOVA, with distribution (novel vs. familiar tone distribution, within subjects), condition (experimental vs. control, between subjects), habituation (habituation to distribution A vs. B, between subjects), and test order (novel vs. familiar distribution first) revealed a significant interaction between distribution and condition, $F\ (1, 40) = 5.213$, $p < .05$. There were no other significant main effects or interactions. Figure 1.10 shows that infants in the experimental group looked longer during presentation of post-habituation melodies having a novel versus familiar tone distribution, while infants in the isochronous control condition showed no preference. Post-hoc Bonferroni-corrected t-tests confirmed a significant novelty preference for infants in the experimental condition, $t\ (23) = 2.44$, $p < .025$, but no preference in the control condition, $t\ (23) = .91$, n.s.

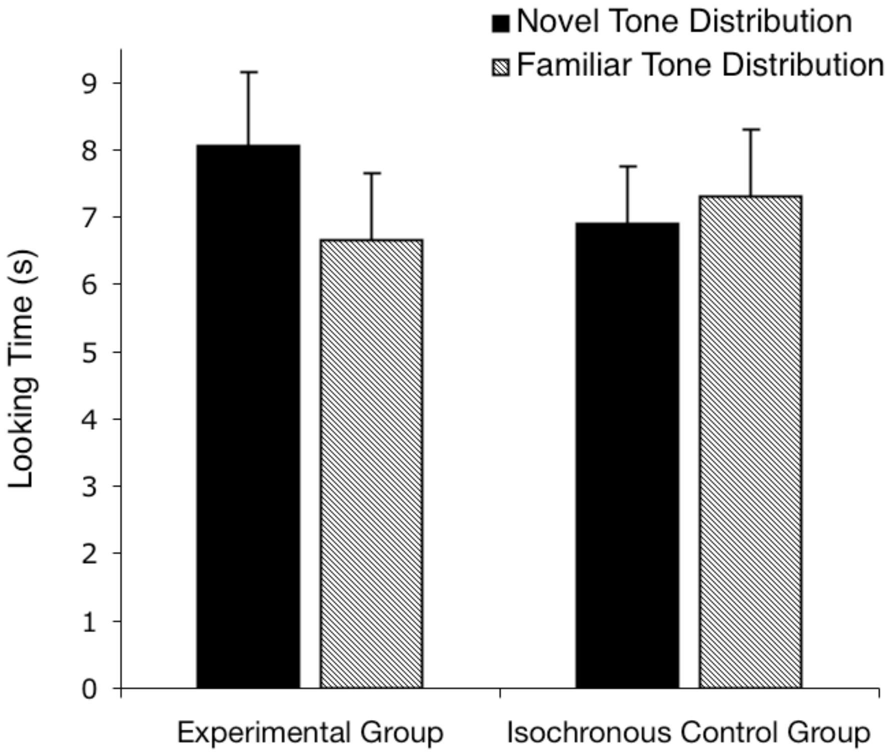


Figure 1.10. Mean looking times after habituation (in seconds). Error bars indicate standard errors. Infants in the experimental group looked longer during presentation of melodies having a novel tone distribution than a familiar tone distribution, but infants in the isochronous control condition showed no preference.

This result indicates that infants in the experimental condition learned an association between the metrical structure and the pitch events. We propose that infants learned during habituation that certain pitches were more likely to occur at some metrical positions and not others. This allowed them to differentiate between two post-habituation melodies having novel or familiar pitch distribution properties. Because infants in the control condition were presented with melodies that were identical in sequential pitch structure but lacking only metrical structure, the absence of a preference in the control condition strongly suggests that the interaction between metrical and pitch structure was responsible for the preference observed in the experimental condition. To our knowledge, these are the first results showing that meter can serve as a framework for infants' learning about properties of pitch structure.

1.5. General Discussion

The present experiments illuminate one way that infants might utilize distributional regularities in auditory input to make sense of complex, rapidly changing auditory patterns such as music. The first two experiments showed that 7-month-old infants categorized unique rhythmic patterns on the basis of underlying metrical structure. Because the rhythms only differed in the frequency of event and accent occurrences at regular periodic positions in the pattern, we can conclude that like adults, infants inferred the meter from these regularities. Experiment 2 further supported this conclusion by pitting two types of structure against each other and showing that implied meter and not grouping structure drove infant preferences. These two experiments provide the most direct evidence to date that infants can infer the underlying meter in rhythmic patterns.

The third experiment revealed one potential function of meter in infancy. Very little is known about how adults acquire knowledge of the complex and hierarchical organization of musical pitch, but some evidence suggests that meter may highlight this structure (Jarvinen & Toiviainen, 2000; Palmer & Pfordresher, 2003). We demonstrated that infants are sensitive to contingencies between pitch events and positions within a metrical framework by showing that they responded differentially to sequences in which those contingencies were reversed. The same pitch sequences without metrical structure failed to elicit such a preference, further supporting the notion that associations between pitch and meter were crucial for infant responding. Although these findings suggest that infants can identify associations between pitch and meter, questions remain about whether infants and children can learn about the relative prominence of a pitch from its frequent placement at strong vs. weak positions in the meter. Because adult-like knowledge of tonality in music has not been observed until childhood (Krumhansl & Keil, 1982; Trehub et al., 1986; Trainor & Trehub, 1992, 1994), it is likely that such learning takes place over a period of years. Other types of information likely contribute to structure learning in music, such as individual pitch frequencies (Creel & Newport, 2002; Krumhansl, 1990). A challenge for future research is to examine whether associations between meter and pitch can lead to adult-like perceptions of structural relationships in music.

Overall, our findings indicate that meter is a highly salient structure in the perception of temporal patterns, even for infants. It is currently unknown whether younger infants can perceive meter. It is possible that meter is learned prior to 7 months, through pre- and post-natal exposure to rhythmic patterns and music (Sansavini, 1997). Certainly many aspects of metrical perception are shaped by experience. For example, unlike adults, 6-month-old infants can detect temporal alterations to musical patterns regardless of metrical conventionality (see Chapter II). This indicates that some culture-specific biases in perceiving meter must be learned between 6 months of age and adulthood. Developmental changes have been observed in synchronized tapping to rhythmic patterns and music, with individuals tapping at an increasingly wide range of slow and fast metrical levels from age 4 to adulthood (Drake et al., 2000). Other studies have documented improvement from age 7 to age 9 in classification and discrimination of metrical versus nonmetrical rhythms (Wilson, Wales, & Pattison, 1997). Individuals may develop more complex temporal representations of meter with a greater number of hierarchical levels as they become increasingly familiar with the musical structures typical of their culture and as their ability to attend to slower metrical levels increases. It is also possible that certain basic aspects of metrical perception arise from a fundamental drive towards synchronization that characterizes the behavior of animate and inanimate systems alike (Strogatz, 2003). Although 7-month-old infants cannot yet coordinate their movements precisely in time, it has been postulated that attentional rhythms can become entrained to external events (Jones & Boltz, 1989; Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999). Infants may easily pick up on periodic regularities in temporal structures they encounter, which makes such regularities a powerful source of information for learning non-temporal information as well as the more complex aspects of temporal structure found in music.

The present findings support a potential function of meter in learning pitch structure, but meter may have other general functions as well. Temporal regularity appears to enhance performance on learning and memory tasks. For example, a regular underlying beat has been shown to improve adults' ability to recall and reproduce rhythmic patterns (Povel & Essens, 1985) and to detect pitch changes (Jones et al., 2002). Highly rhythmic, metrical music facilitates performance on standardized spatial-temporal tasks (Thompson, Schellenberg, & Husain, 2001) and recall of autobiographical details by elderly adults suffering from dementia (Foster & Valentine, 2001). Periodic regularity in speech may also enhance memorization in special cases such as poetry or other oral traditions (Rubin, 1995). Scholars have hypothesized that temporal regularity in caregiver singing and speech may play a fundamental role in regulating infant arousal and enhancing infant learning of linguistic, musical and social information (Stern, Spieker, Barnett, & MacKain, 1983; Trainor et al., 1997). Meter, as one manifestation of periodic temporal structure, may provide important benefits for learning and remembering auditory patterns in general.

A growing body of evidence suggests that general learning mechanisms likely enable infants to detect and utilize distributional regularities for parsing and interpreting complex sequential input. Many of these regularities depend on temporal information, such as simultaneity (e.g., Mattyz & Jusczyk, 2001) or temporal proximity (e.g., Kirkham, Slemmer, & Johnson, 2002; Saffran et al., 1996; Thiessen & Saffran, 2003). Infants can even learn to associate units that are temporally non-adjacent (Gomez & Gerken, 2002; Newport & Aslin, 2004). Our findings are the first to demonstrate that infants can detect temporal regularities that occur periodically. While periodic temporal structure may play only a relatively minor role in the alternating stress patterns of speech, it is fundamental to music perception and behavior. Meter may thus function as a tool for bootstrapping knowledge about the organization of music, without which individuals could not participate in common musical activities such as listening, dancing, performing, or remembering a familiar tune.

CHAPTER II.

Metrical Categories in Infancy and Adulthood*

2.1. Introduction

The most important aspect of music may be its capacity to facilitate human movement coordination. All cultures have sound patterns with repetitive temporal structures, which facilitate synchronous dancing, clapping, playing instruments, marching, and chanting (Brown, 2003). These communal activities imply universal propensities to coordinate movement in time. On occasion, however, listeners are challenged by the rhythmic structures of foreign music. For example, North American adults have difficulty perceiving the rhythmic organization of classical Indian music despite the ease with which Indian audiences clap in time. Such phenomena implicate processes of musical enculturation in shaping perceptual abilities.

Many scholars have documented powerful biases in the perception and production of rhythmic patterns. Musical patterns contain a range of interonset interval (IOI) durations, but Western adults often fail to respond differentially to small but perceptible duration differences. Instead, they seem biased to perceive and produce durations with simple ratios. For example, their spontaneous productions reveal long and short durations in a 2:1 ratio (Fraisse, 1982), and their reproductions of rhythmic patterns reveal durations stretched or reduced to fit simple ratios (Collier & Wright, 1995; Cummins & Port, 1998). Even musicians mistakenly assign simple-duration ratios to rhythms characterized by complex ratios (Desain & Honing, 2003). Once listeners assign a specific duration ratio, they continue to interpret the unfolding pattern in terms of that ratio despite perceptible temporal changes (Clarke, 1987;

---

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

Large, 2000). These findings are consistent with the categorization of duration ratios and with the assimilation of complex- to simple-duration ratios.

Biases for simple duration ratios are thought to arise from adults' categorization of durations according to the hierarchical temporal structure of the music, or the *meter*. On hearing a rhythmic musical pattern, listeners tend to infer a primary, underlying pulse of equally spaced (i.e., isochronous) events; faster levels of this underlying pulse result from binary or ternary subdivisions. Meter gives rise to the perception of alternating strong and weak beats, as in the waltz pattern of "**one** two three, **one** two three…" which results from the convergence of faster and slower isochronous levels in a 3:1 ratio (see Figure 2.1). Theoretical accounts of meter assume that the durations of a specific rhythmic pattern are assimilated to an internal periodic clock (Povel, 1981; Povel & Essens, 1985) or system of oscillators (Large & Kolen, 1994). Rhythms containing simple duration ratios are thought to have greater coherence because they are readily assimilated to such isochronous, hierarchical structures (Jones & Boltz, 1989; Jones, Moynihan, MacKenzie, & Puente, 2002).

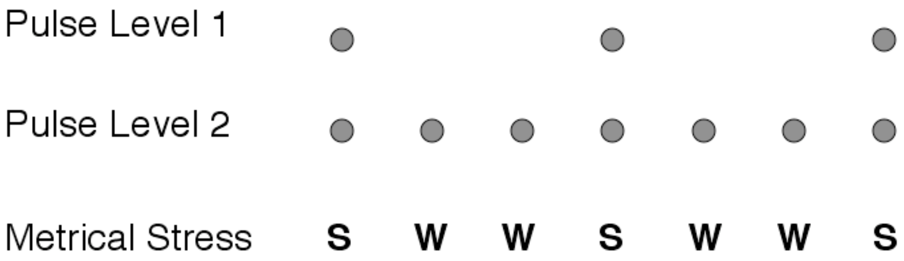


Figure 2.1. Alternating strong (S) and weak (W) events in a metrical hierarchy of two isochronous pulse levels having a 3:1 ratio, typical of a waltz.

On the basis of the aforementioned findings, many scholars contend that intrinsic perceptual biases constrain the organization of rhythmic patterns.

Specifically, “good” metrical structures are thought to consist of isochronous patterns at nested hierarchical levels that are related by simple integer ratios (Clarke, 1999; Povel & Essens, 1981). An alternative explanation is that the perceptual biases of Western listeners arise from the predominance of simple ratios in Western music. This possibility has not been evaluated experimentally despite the occurrence of metrical structures that violate the foregoing assumptions (isochronous levels and simple-integer ratios) in many non-Western musical cultures.

Much of the music from Eastern Europe, South Asia, the Middle East, and Africa has an underlying pulse of alternating long and short durations in a 3:2 ratio (Clayton, 2000). These complex meters, which are common in Bulgarian and Macedonian folk music, pose no apparent problems for adult and child singers and dancers from those countries (London, 1995; Rice, 1994). Prolonged exposure to specific metrical structures may enable listeners to distinguish structurally meaningful timing changes from changes (e.g., slowing down) associated with expressive performance (Desain & Honing, 2003). Instead of interpreting and reinterpreting unfolding sequences on the basis of small but detectable temporal deviations, listeners interpret these sequences within the framework of the meter (Clarke, 1999). Presumably, implicit knowledge of metrical structure, which undoubtedly varies across cultures, is central to the perception of rhythmic patterns.

Because infants have limited experience with music, they provide unique opportunities for examining intrinsic biases for metrical structure. In principle, comparisons between infant and adult listeners could reveal biases that stem from musical enculturation or from perceptual predispositions. In the speech domain, adults typically perceive native consonant contrasts more readily than foreign contrasts because of their tendency to assimilate foreign sounds to the perceptual categories of their native language (Best, McRoberts, & Sithole, 1988). By contrast, infants’ perception of native and foreign consonant contrasts is equivalent until about 10-12 months of age, when native consonant categories begin to emerge (Werker & Lalonde, 1988). In the musical domain, analogous initial abilities and subsequent assimilation processes lead infants to exhibit comparable pitch discrimination skills in native and foreign melodic contexts and adults to exhibit superior skills in native than in foreign melodic contexts (Lynch, Eilers, Oller, & Urbano, 1990; Trehub, Schellenberg, & Kamenetsky, 1999). The implication is that exposure to speech and music leads to culture-specific fine-tuning or perceptual reorganization.

If infants begin life with relatively flexible perceptions of meter, they should process simple- and complex-duration ratios equally well. By contrast, human predispositions for simple-duration ratios, like those for simple frequency ratios (Schellenberg & Trehub, 1996), would lead infants to perceive simple-duration ratios more efficiently than complex-duration ratios. Regardless of the presence or absence of infant biases, adult attunement to the metrical categories of their musical culture should lead to enhanced processing of culturally typical duration ratios. In the present study, we compared North American adults, whose exposure was largely restricted to simple-duration ratios, Bulgarian or Macedonian adults, whose exposure included both simple- and complex-duration ratios, and 6-month-old infants, whose limited musical exposure involved simple-duration ratios. Specifically, we evaluated their perception of temporal changes in folk melodies with simple or complex metrical structure.

2.2. Experiment 1

After familiarization with multi-instrument performances of foreign folk melodies, North American adults were presented with altered versions that preserved or violated the original metrical structure. Listeners rated the consistency of the alterations with the rhythmic structure of the original performance.

2.2.1. Method

2.2.1.1. Participants

Participants were 50 college students (35 women, 15 men, 18-25 years) whose musical training ranged from 0-15 years (*M* = 5.7). Those with 7+ years of music lessons were considered highly trained. Most participants had lived exclusively in North America, but three had lived in England, Russia, or the Dominican Republic for 6-10 years.

2.2.1.2. Apparatus and Stimuli

Four excerpts were taken from traditional folk-dance melodies of Serbia and Bulgaria (Geisler, 1989). All excerpts consisted of eight cycles, or measures, of seven or eight notes per measure in the complex- or simple-meter excerpts, respectively. Each excerpt was arranged as a MIDI performance with four Quicktime MIDI Instruments. The instrumentation consisted of a primary melodic instrument (*Acoustic Fretless Bass* or *Tango Accordion*), a secondary melodic instrument (*Flute*), an accompanying harmonic instrument (*Acoustic Fretless Bass* or *Bright Acoustic Piano*), and a percussion instrument (*Melodic Tom*, *Timpani*, or *Kalimba*). Two excerpts were from dances in simple meter (each eight-note measure subdivided into groups of 2+2+2+2), and two were in complex meter (each seven-note measure subdivided into groups of 2+2+3 or 3+2+2, depending on the type of dance). Both primary melodic instruments were used in each meter.

Most durations, or interonset intervals (IOIs), in the primary melodic line were 250 ms, but longer IOIs (500, 750, and 1000 ms) occurred in all excerpts, primarily at strong metrical positions. To highlight the primary metrical pulse, note amplitude (MIDI velocity) was increased at strong metrical positions (velocities of 120-127 relative to velocity of 90 for all other notes). Drum accompaniment and harmonic instruments also occurred at strong metrical positions. Thus, the familiarization stimuli provided a variety of cues that are considered important for inferring the meter of a sequence.

Test stimuli consisted of altered versions of the familiarization melodies presented at uniform amplitude (MIDI velocity of 90). These stimuli were simplified by the use of one melodic instrument (*Acoustic Grand Piano*) and one drum accompaniment (*Woodblock*). For all stimuli, the drum accompaniment consisted of patterns of long and short durations that were repeated in every measure. For simple- and complex-meter stimuli, this accompaniment consisted of Long-Short-Short or Short-Short-Long patterns.

Test melodies were identical to the primary familiarization melody except for a structure-preserving or structure-violating alteration. Familiarization, structure-preserving, and structure-violating stimuli for simple and complex meters are depicted musically and graphically in Figure 2.2. (Audio excerpts are available at http://people.psych.cornell.edu/~eeh5/MCstimuli.html).

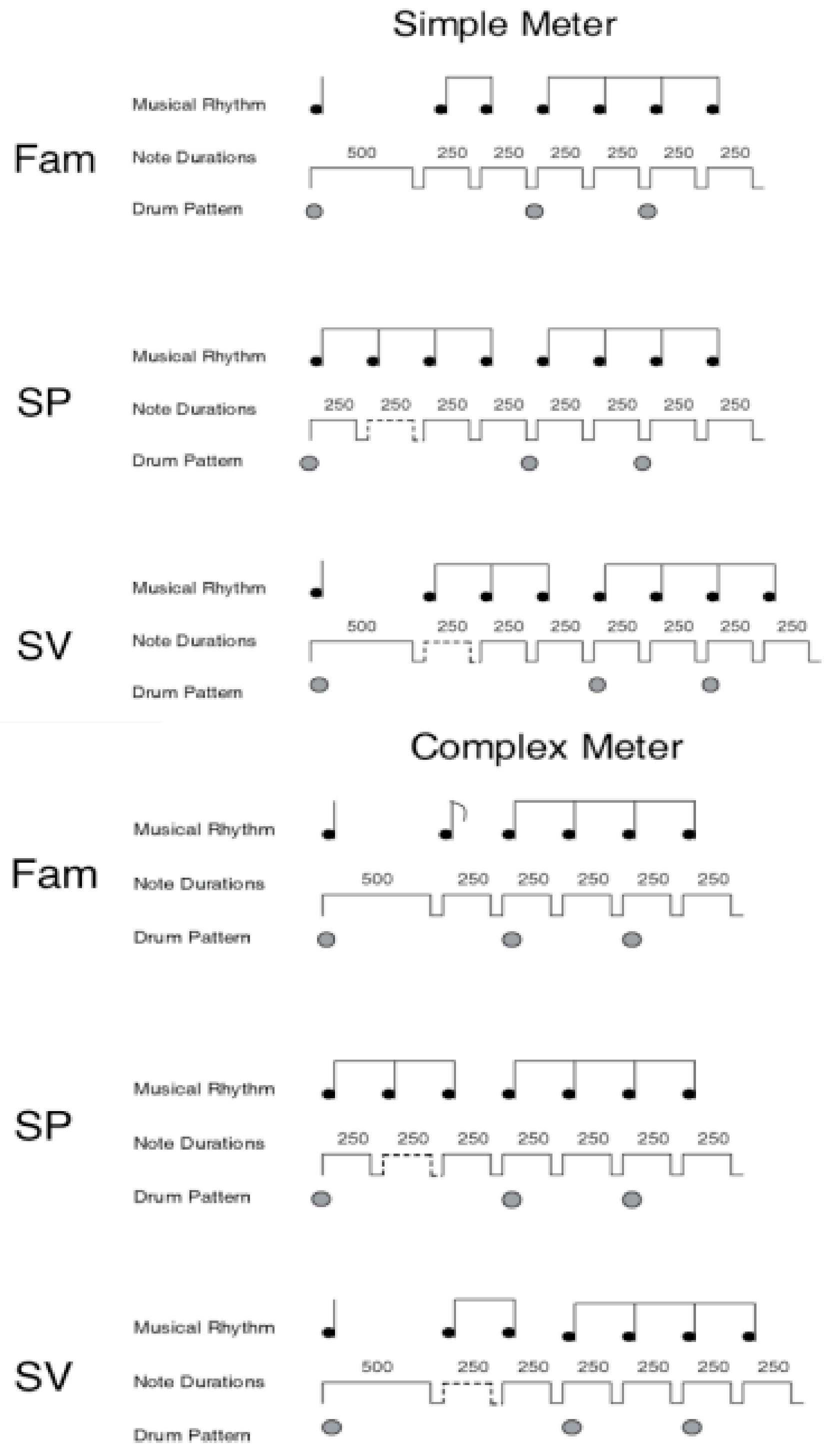


Figure 2.2. Displayed in musical notation and graphical form are one-measure examples of a typical familiarization stimulus (Fam), its structure-preserving alteration (SP) and its structure-violating alteration (SV), for both simple and complex meter.

The drum accompaniment is indicated by the gray dots. For test stimuli the added note is indicated by a dashed line.

A single note, identical in pitch and location, was inserted in both structure-preserving and structure-violating alterations. Only in structure-preserving versions were the durations of adjacent notes modified to preserve the original metrical structure. Figure 2.2 indicates the alteration of the rhythmic structure of each stimulus by the insertion of a 250-ms note (dashed line) after the first note of the measure. For structure-preserving alterations, the preceding note duration was reduced from 500 to 250 ms, which preserved the overall duration of the measure and the pattern of durations in the drum accompaniment. The structure-preserving drum accompaniment consisted of long and short durations of 1000 and 500 ms for the simple meter and 750 and 500 ms for the complex meter. Structure-violating alterations had no adjustment to adjacent notes, leading to an increase in the long duration of the drum pattern and in the overall duration of the measure. Thus, structure-violating drum accompaniments consisted of long and short durations of 1250 and 500 ms for the simple meter and 1000 and 500 ms for the complex meter. Note that structure-violating alterations of the simple meter resulted in a complex meter and structure-violating alterations of the complex meter resulted in a simple meter.

Each measure of each stimulus had an alteration that varied in location but always occurred during the long duration of the drum accompaniment. To minimize the salience of the melodic change, the pitch of the inserted note was identical to the preceding or following pitch. Because the inserted note was identical in pitch and location for both alteration types, structure-preserving and structure-violating test stimuli were identical except for the ratio change between short and long durations of the drum pattern and the overall duration of the measure. Aside from the structure-preserving and structure-violating stimuli, two additional test stimuli served as foils. One was an unaltered version of the primary familiarization melody, and the other was disrupted by the pseudo-random insertion of notes twice per measure. These foils were included to promote a wide range of ratings.

We generated one familiarization stimulus and four test stimuli for the practice trials that preceded testing. The familiarization stimulus was a well-known children's tune (*Mary Had a Little Lamb*). Two test stimuli preserved the rhythmic structure, one being identical to the familiarization stimulus, the other having extra notes that preserved the rhythm. Two test stimuli had obvious violations of rhythmic structure resulting from the addition of two or three notes or pauses per measure. All participants rated these structure-violating test stimuli as inconsistent with the familiarization stimulus.

2.2.1.3. Procedure

Participants, who were tested in groups of 1-5, listened to the stimuli over headphones at individual computer stations. All musical excerpts and instructions were controlled by means of PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Trials were presented in blocks consisting of one familiarization stimulus followed by four test stimuli: two structure-preserving and two structure-violating versions of the familiarization stimulus. Participants listened to a 2-min familiarization sequence, consisting of seven repetitions of simple-meter excerpts and eight repetitions of complex-meter excerpts. Then they rated how well the four test stimuli matched the rhythmic structure of the familiarization stimulus (1, or very well, to 6, or very poor). Each block was repeated three times, resulting in three sets of judgments per stimulus. Order of blocks and order of test stimuli were counterbalanced across participants.

2.2.2. Results and Discussion

A three-way mixed-design ANOVA, with familiarization meter (simple vs. complex familiarization stimulus, within subjects), alteration type (structure-preserving vs. structure-violating, within subjects), and musical training (high vs. low, between subjects), revealed significant main effects of meter, $F(1,48) = 18.18$, $p < .001$, $\eta^2 = .28$, and alteration type, $F(1,48) = 88.59$, $p < .001$, $\eta^2 = .65$, and a significant interaction between meter and alteration type, $F(1,48) = 117.12$, $p < .001$, $\eta^2 = .71$. Inspection of Figure 2.3 shows that much higher ratings (i.e., greater dissimilarity) were given to structure-violating than to structure-preserving alterations for simple-meter excerpts but not for complex-meter excerpts. The interaction between musical training, meter, and alteration type was not significant, $F(1,48) = 2.06$, $p = .16$, $\eta^2 = .04$, but there were indications that musical training affected response accuracy. Accuracy scores were calculated by subtracting ratings of structure-preserving alterations from ratings of structure-violating alterations. Accuracy scores and years of musical training were correlated for simple meter stimuli, $r(48) = .43$, $p < .01$, but not for complex meter stimuli, $r(48) = .18$, $p = .21$. In other words, more extensive musical training was associated with more differentiated responding to structure-violating and structure-preserving stimuli only in the context of familiar metrical structure.

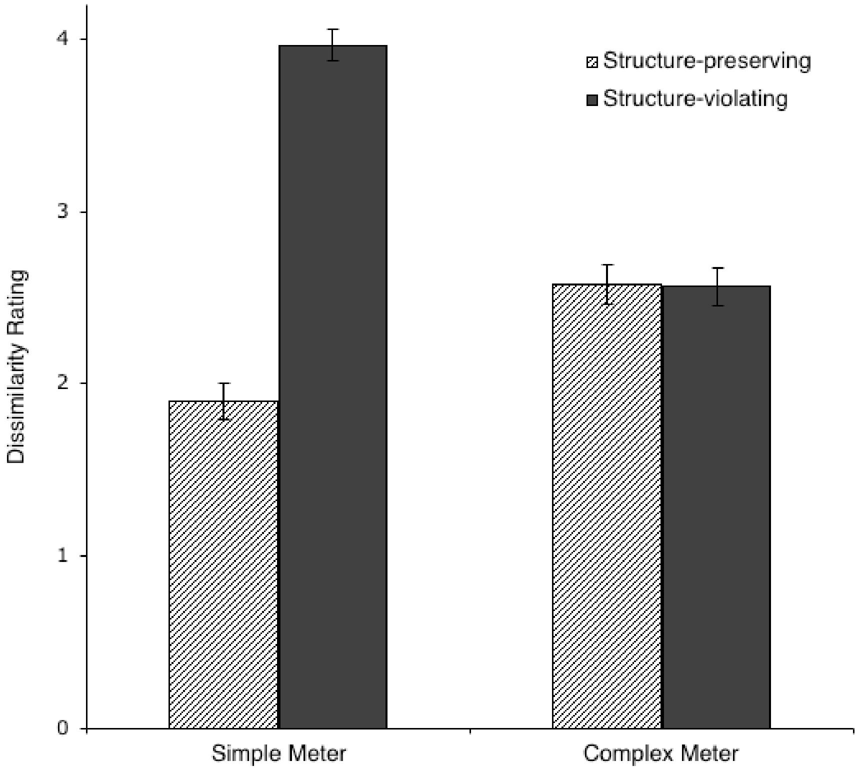


Figure 2.3. North American adults' mean dissimilarity judgments for structure-preserving and structure-violating alterations to simple- and complex-meter patterns (Experiment 1). Error bars indicate standard errors.

In summary, the metrical structure of musical patterns had a dramatic effect on the perception of those patterns. For simple metrical structures, adult ratings of alterations that disrupted simple-duration ratios and underlying isochronous structure were significantly different from ratings of alterations that preserved the structure. No such differentiation was evident in the context of complex metrical structure. In effect, adults noticed alterations from typical to atypical metrical structures but not alterations from atypical to typical metrical structures. Such asymmetric performance parallels adults' greater ease of detecting changes from conventional to unconventional pitch and rhythmic patterns than from unconventional to conventional patterns (Bharucha &

Pryor, 1986; Schellenberg, 2002). Encoding difficulties with unconventional sequences may arise from the inappropriate assimilation of atypical sequences to familiar musical categories.

2.3. Experiment 2

Adults' difficulty with complex metrical structures in Experiment 1 is consistent with musical enculturation processes. Nevertheless, it remains to be determined whether adults exposed to simple and complex ratios, like those from Bulgaria or Macedonia (Rice, 1994), would exhibit comparable ease of processing for complex- as well as simple-duration ratios. The present experiment was a replication of Experiment 1 with adults of Bulgarian or Macedonian origin.

2.3.1. Method

2.3.1.1. Participants

Participants were 17 first- or second-generation immigrants (12 women, 5 men, 18-38 years) from Bulgaria (n = 13) or Macedonia (n = 4) who had 0-14 years (M = 3.4) of music lessons (primarily involving Western music). All of them had participated in traditional cultural activities in childhood and/or adulthood, either in their home country or in the host country.

2.3.1.2. Apparatus and Procedure

The apparatus, stimuli, and procedure were identical to Experiment 1.

2.3.2. Results and Discussion

A three-way mixed-design ANOVA, with familiarization meter (simple vs. complex, within subjects), alteration type (structure-preserving vs. structure-violating, within subjects), and musical training (high vs. low, between subjects), revealed a significant main effect of alteration type, $F(1,15) = 49.33$, $p < .001$, $\eta^2 = .77$, and a significant interaction between alteration type and musical training, $F(1,15) = 11.69$, $p < .01$, $\eta^2 =$

.44. There was no main effect of familiarization meter and no other significant interactions. Inspection of Figure 2.4 reveals higher ratings (greater dissimilarity) of structure-violating than structure-preserving alterations for both simple- and complex-meter excerpts.

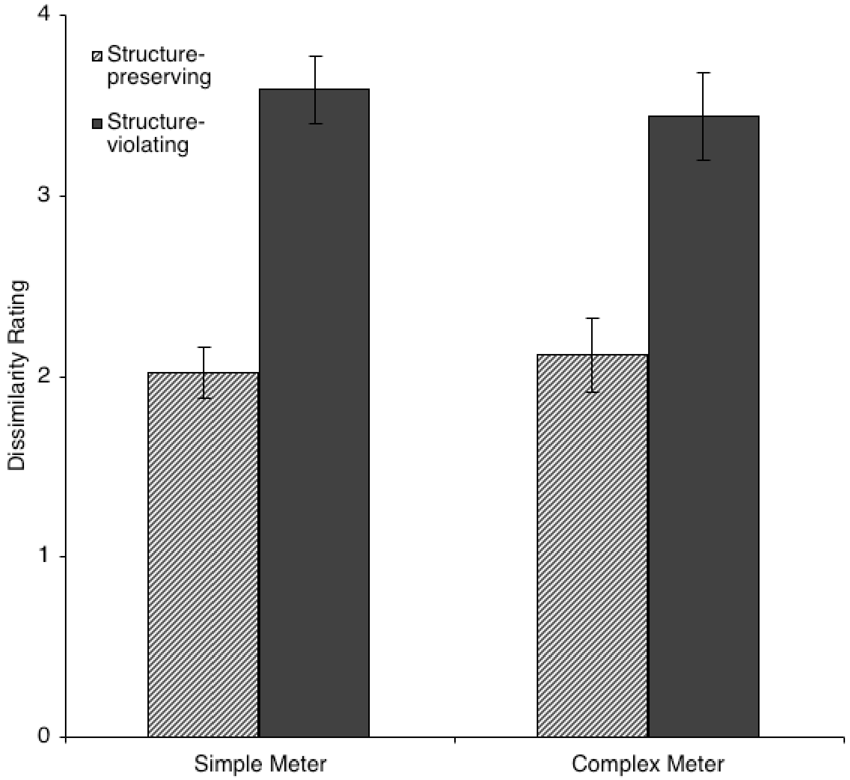


Figure 2.4. Bulgarian and Macedonian adults' mean dissimilarity judgments for structure-preserving and structure-violating alterations to simple- and complex-meter patterns (Experiment 2). Error bars indicate standard errors.

Accuracy scores, calculated as in Experiment 1, were significantly correlated with years of musical training for simple-meter stimuli, $r(15) = .55$, $p < .05$, and for complex-meter stimuli, $r(15) = .60$, $p < .05$. Unlike adults of North American origin, those of Bulgarian and Macedonian origin had experienced long-term exposure to simple and complex metrical structures, which enabled them to differentiate structure-violating alterations from structure-preserving alterations in both metrical contexts.

2.4 Experiment 3

Although the findings of Experiments 1 and 2 were consistent with enculturation processes, they did not rule out inherent biases for simple metrical structures. We investigated 6-month-old infants' perception of the simple- and complex-meter stimuli of Experiments 1 and 2 by means of a familiarization-preference procedure.

2.4.1. Method

2.4.1.1. Participants

Participants were 64 infants who were 6-7 months of age (*M* = 6.9 months; 28 girls, 36 boys). All infants were free of colds on the test day and had no family history of hearing impairment. An additional 21 infants were excluded from the final sample because of fussing (*n* = 18), sleeping (*n* = 1), or technical failure (*n* = 2).

2.4.1.2. Apparatus and Stimuli

Familiarization and test stimuli consisted of the four sets of simple- and complex-meter stimuli from Experiment 1. Half of the infants received one of the two simple-meter stimuli; the other half received one of the two complex-meter stimuli. All stimuli were prepared as Quicktime movies accompanied by identical visual (non-rhythmic) portions of a documentary film (Attenborough, 1991).

A Macintosh G4 computer controlled (a) the presentation of auditory stimuli though a centrally located but out-of-view loudspeaker (Altec Lansing AC 522), (b) the presentation of visual stimuli on two 17-in color monitors (Sony) separated by approximately 91 cm, and (c) the recording of infant responses. The experimenter, who observed the infant through a small hole in a partition separating the infant from the experimenter and the control equipment, recorded all changes in infants' direction of gaze. Infant visual fixations were also visible from the monitor of a digital video camera that focused on the infant's face. The experimenter and mother wore headsets playing music to mask the auditory stimuli presented to infants.

2.4.1.3. Procedure

We tested infants by means of a familiarization-preference procedure. Infants sat on their parents' lap in a dimly lit testing room. The two monitors were approximately 140 cm in front and to the right and left of the seated infant. The observer recorded infant looking times by pressing one of two buttons on the computer, one button for looking towards a monitor, the other for looking away. Infants were first presented with 2 min of the familiarization stimulus, consisting of four 30-s repetitions of the stimulus that alternated between monitors. To attract infants' attention to the appropriate monitor, a flashing red screen on that monitor preceded each trial. When the infant looked at the monitor, a familiarization trial began and continued for 30 s. After four familiarization trials, test stimuli were presented six times each, with the structure-preserving and structure-violating test stimuli alternating between monitors. Each test trial was terminated when the infant looked away for 2 s or when 60 s had elapsed.

Infants were randomly assigned to one of the two familiarization excerpts in simple or complex meter. The order of the first monitor in the familiarization phase, the first monitor in the test phase, and the first test stimulus (structure-preserving vs. structure-violating) was counterbalanced.

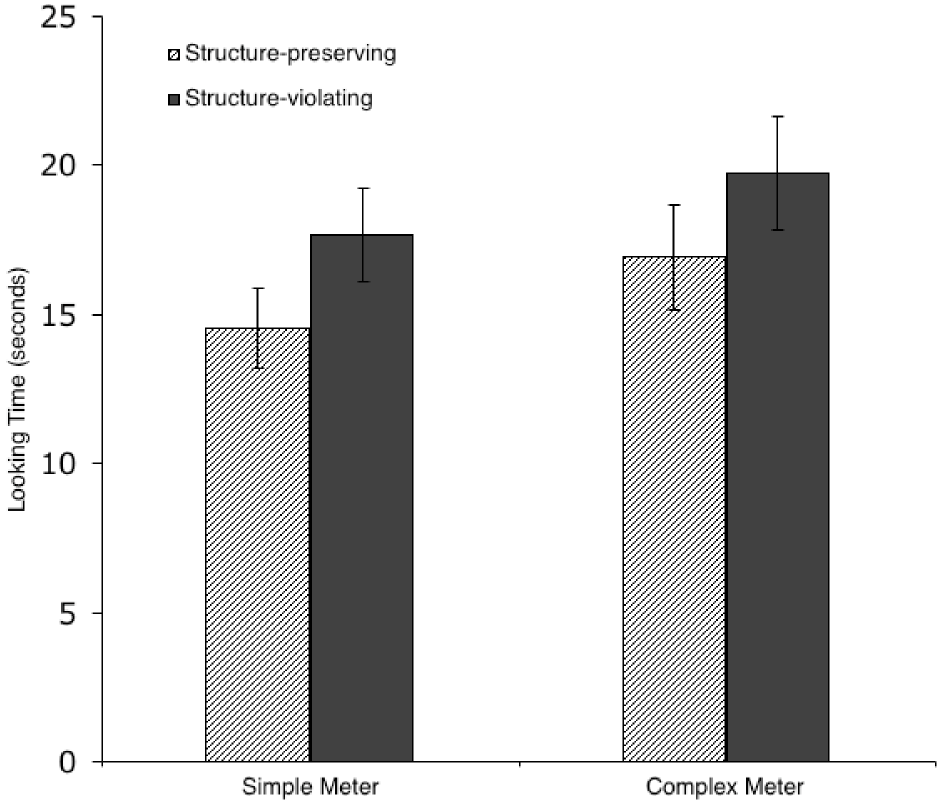


Figure 2.5. Infants' mean looking times (in seconds) for structure-preserving and structure-violating alterations to simple- and complex-meter patterns. Error bars indicate standard errors.

2.4.2. Results and Discussion

Looking times were analyzed for differential attention to structure-preserving and structure-violating alterations in the conditions with simple and complex meter. A two-way, mixed design ANOVA, with familiarization meter (simple vs. complex, between subjects) and alteration type (structure-preserving vs. structure-violating, within subjects) revealed a significant main effect of alteration type, $F(1, 62) = 16.445$, $p < .001$, $\eta^2 = .21$. There was no significant interaction between familiarization meter and alteration type. Figure 2.5 presents mean looking times for structure-preserving and structure-violating test stimuli in the simple- and complex-meter conditions. In the context of simple and complex meters, infants looked significantly longer for structure-violating alterations than for structure-preserving alterations. Greater looking time to structure-violating alterations is interpretable as a novelty preference, which implies that infants perceived structure-violating alterations as less similar to the familiarization stimulus than structure-preserving alterations. Unlike North American adults, who differentiated structure-violating from structure-preserving changes only in the context of simple-meter excerpts, infants differentiated those changes in complex- as well as simple-meter contexts.

2.5. General Discussion

North American infants were more similar to Bulgarian and Macedonian adults than to North American adults because they differentiated alterations that violated the metrical structure from those that preserved the metrical structure of musical patterns with simple or complex meter. These findings imply that the temporal perception and production biases of North American adults arise from extended exposure to the simple metrical structures that predominate in Western music. Human listeners may begin with flexible processing of metrical structure, which facilitates perception of temporal nuances in various kinds of music. Months, perhaps years, of exposure to the dominant metrical categories of a specific musical culture may prompt perceptual reorganization or narrowing of the metrical frameworks that can be handled with ease.

It is possible, although unlikely, that the absence of infant biases for simple ratios reflects different mechanisms for processing temporal structures in infancy and adulthood (Collier & Wright, 1995). For example, infants could perceive temporal patterns in a serial manner, without the hierarchical and anticipatory aspects of adult metrical processing. However, several lines of evidence are consistent with the idea of rich metrical processing in infancy. Infants have a number of prerequisite skills for perceiving meter, including the detection of subtle changes in duration (Morrongiello & Trehub, 1987) and tempo (Baruch & Drake, 1997), the detection of miniscule (12 ms) gaps in brief tones (Trehub, Schneider, & Henderson, 1995), the discrimination of isochronous from non-isochronous tone patterns (Demany, McKenzie, & Vurpillot, 1977), and the generalization of auditory patterns on the basis of rhythmic structure (Trehub & Thorpe, 1989). Infants can discriminate musical excerpts on the basis of rhythmic changes (Chang & Trehub, 1977) and on the basis of subtle performance cues (e.g., intensity and duration changes) associated with metrical structure (Palmer, Jungers, & Jusczyk, 2001). These cues are correlated with metrical structure in infant-directed singing (Trainor, Clark, Huntley, & Adams, 1997). Moreover, infants' ability to detect small timing changes depends on the strength of implied metrical structure (Bergeson, 2002). Finally, 7-month-old infants can categorize unique rhythms on the basis of implied metrical structure (see Chapter I). In view of the aforementioned evidence, it is likely that meter is fundamental to the organization of auditory-temporal input in infancy. Greater flexibility in that organization may be possible early in life, before listeners become attuned to the musical input in their environment.

Our findings imply that adult biases in temporal pattern processing result from category learning processes that are part of musical enculturation, not from intrinsic perceptual biases for simple temporal structures. Implicit knowledge of musically relevant categories is critical for the appreciation of music in any culture. Listeners must discern the metrical structure of a piece in the face of temporal fluctuations that reflect the performer's expressive intentions (Desain & Honing, 2003). Comparable category learning enables listeners to discern words and meaningful prosodic changes in the speech stream despite enormous variability within and across speakers.

Abilities that are part of the *initial state* of auditory pattern processing are likely to undergo reorganization when young listeners discover which distinctions are common or meaningful in their culture and which are not. For some segmental and suprasegmental aspects of speech, perceptual reorganization occurs in infancy (Nazzi, Jusczyk, & Johnson, 2000; Werker & Lalonde, 1988). Comparable changes in music processing may have a more protracted course of development. For example, culture-specific changes in the perception of musical harmony do not occur until the early school years (Krumhansl & Keil, 1982; Trainor & Trehub, 1994). Our findings provide the first demonstration of reorganization in temporal pattern processing, presumably as a result of exposure to music. One challenge for the future is to document the developmental course of that reorganization.

CHAPTER III.

Tuning in to Musical Rhythms: Infants Learn More Readily than Adults

3.1. Introduction

The ability to recognize and respond appropriately to species-specific information is essential for communication and survival. Experience-dependent narrowing, or "tuning" in the first year after birth may facilitate the acquisition of perceptual skills in a range of socially meaningful domains. The domain of speech is a prominent example. Initially, infants discriminate speech sounds from languages they have never heard, but over the first year they become differentially responsive to a narrower range of speech distinctions that are relevant only in their native language-to-be (Werker & Tees, 1984; Kuhl, Williams, Lacerda, & Stevens, 1992). A similar developmental course is evident in the domain of face perception. For example, 6-month-olds differentiate individual faces of nonhuman as well as human primates, but 9-month-olds are more like adults in differentiating human faces only (Pascalis, deHaan, & Nelson, 2002). Here we show comparable experience-dependent tuning in the domain of musical rhythm perception. We also demonstrate that culture-specific musical biases, once acquired, are more resistant to change in adulthood than in infancy.

Synchronized movement to music, such as clapping, tapping, dancing, singing, and ensemble performance, has been observed across all known cultures and historical periods, which implies universality of this aspect of human behavior (Brown, 2003). Such behavior is thought to depend on the ability to infer an underlying musical "beat" or *meter,* and to integrate rhythmic information into that metrical framework[2]. In

---

[2] *Meter* is the underlying pattern of strong and weak beats inferred from the musical surface, as in the waltz pattern of "*one* two three, *one* two three…." In Western music, meter consists of multiple hierarchical levels of evenly spaced (i.e., isochronous) periodic structure, with faster levels resulting

perception and production tasks, adults exhibit a powerful tendency to assimilate continuously varying rhythmic information into a familiar (i.e., culture-specific) metrical framework. Because Western listeners are accustomed to the temporally even or “isochronous” meters of Western music, they may have considerable difficulty remembering or reproducing patterns that are not isochronous. Even when target patterns have noticeable deviations from isochrony, Western adults stretch or shrink the component rhythmic intervals towards an isochronous framework (Collier & Wright, 1995; Desain & Honing, 2003; Essens & Povel, 1985; Fraisse, 1982; Povel, 1981). Such culture-specific biases interfere with adults‘ differentiation of rhythmic variations of non-isochronous (foreign) tunes, even though such rhythms pose no difficulty for 6-month-old infants (see Chapter II). Specifically, North American adults readily detect rhythmic variations that disrupt a familiar (Western), isochronous meter, but they fail to notice comparable disruptions of a foreign (Balkan), non-isochronous meter. By contrast, adults from Bulgaria and Macedonia, who receive exposure to both types of meter in childhood, and 6-month-old infants, who receive no such exposure, distinguish rhythmic variations in both isochronous and non-isochronous contexts. This finding implies culture-specific responding to musical rhythms by adult listeners but culture-general responding by 6-month-old listeners, consistent with findings in early speech and face processing (Eimas, Siqueland, Jusczyk, & Vigorito, 1971; Pascalis et al. 2002).

If musical rhythm perception undergoes a process of experience-dependent tuning that parallels speech and face perception, then the transition from culture-

---

from binary or ternary subdivisions. In some types of Balkan music, meter also consists of multiple levels, but some levels are not isochronous and instead consist of alternating short and long temporal intervals. The term *rhythm* refers to the sequence of temporal intervals in a pattern, such as 500-500-250 ms. Because meter tends to constrain the rhythmic structure of music, isochrony in Western meters gives rise to a greater frequency of simple ratios between temporal intervals, such as 1:1 or 2:1. By contrast, Balkan rhythms frequently contain long and short temporal intervals with complex ratios, such as 3:2.

general to culture-specific responding could occur early in life. We investigated this possibility in Experiment 1 by testing Western 12-month-old infants with the familiarization-preference procedure and stimuli used previously with Western 6-month-olds (Chapter II). Experiments 2 and 3 examined whether culture-specific biases in 12-month-olds and adults could be reversed after brief, at-home exposure to foreign, non-isochronous music.

3.2. Experiment 1

Infants were familiarized with a synthesized performance of a Balkan folk tune, followed immediately by two simplified variations of the folk tune, one containing a change that disrupted the meter of the original, and the other containing a change that preserved the original meter. Infants were familiarized with a tune having either an isochronous meter that is common in both Western and Balkan music, or a non-isochronous meter that is common in Balkan (Bulgarian and Macedonian) music but not in conventional Western music. Infants' looking times were examined for differential attention to structure-disrupting and structure-preserving test stimuli.

3.2.1. Method

3.2.1.1. Participants

Participants were 52 infants who were 11-12 months of age (*M* age = 349.5 days, *SD* = 17.5 days) at the time of testing, 26 girls and 26 boys. All participants were healthy, full-term infants who were free of colds on the test day and had no family history of hearing impairment. An additional 14 infants were excluded from the final sample because of fussing (n = 12) or parental interference (n = 2). Most infants lived in a monolingual English-speaking environment. Parents reported that neither they nor their infants had prior exposure to Balkan music.

3.2.1.2. Stimuli

Familiarization and test stimuli consisted of two isochronous and two non-isochronous melodies used in Hannon and Trehub (Chapter II). All excerpts were selected from a collection of traditional folk-dance melodies from the Balkans (Geisler, 1989). Figure 3.1 illustrates the rhythmic structure for one cycle, or *measure*, of isochronous and non-isochronous meter excerpts. Excerpts in isochronous and non-isochronous meter consisted of eight measures in total, with each measure containing a maximum of eight 250 ms-notes in isochronous meter and seven 250 ms-notes in non-isochronous meter (Figure 3.1). Most notes were 250 ms in duration, but longer notes (500, 750, and 1000 ms) occurred occasionally in all excerpts. For all stimuli, each eight-measure excerpt was cycled repeatedly up to a maximum of approximately one minute (four repetitions).

For familiarization stimuli, each excerpt was computer-generated as a multi-instrument performance, using a MIDI sequencer and four Quicktime synthesized instruments that played melody, harmony, or percussion. The percussion instrument played a drum pattern that subdivided each measure into either a long-short-short or a short-short-long sequence of temporal intervals. For isochronous excerpts, long and short intervals in the drum pattern had a 2:1 ratio. For non-isochronous excerpts, long and short intervals had a 3:2 ratio (see Figure 3.1).

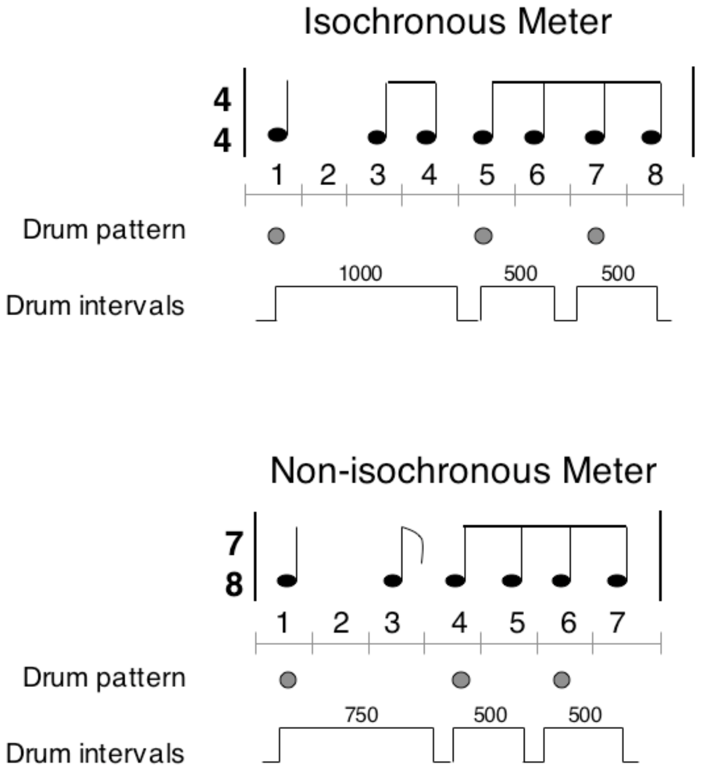


Figure 3.1. One measure each of isochronous and non-isochronous meter familiarization excerpts with a long-short-short drum accompaniment, depicted in musical notation and graphical form. Each count of the measure is numbered to illustrate that isochronous meter excerpts consist of eight counts per measure, while non-isochronous meter excerpts contain seven counts per measure. The intervals in the isochronous meter drum pattern form a long-to-short ratio of 1000:500, or 2:1, while the intervals in the non-isochronous meter drum pattern form a long-to-short ratio of 750:500, or 3:2.

Test stimuli were variations of the familiarization excerpts with reduced complexity, consisting of one melodic instrument and one percussion instrument. Test melodies were identical to familiarization melodies with the exception of a change that either preserved or disrupted the original meter. A single 250-ms note, identical in pitch and location, was inserted in each measure of both structure-preserving and structure-disrupting variations. In structure-preserving variations the durations of adjacent notes were modified to preserve the drum pattern and the number of counts per measure. In structure-disrupting variations the additional note increased the duration of the longer interval in the drum pattern as well as the number of counts per measure (Figure 3.2). Because the extra note was identical in pitch and location for both alteration types, structure-preserving and structure-disrupting test stimuli were identical except for the ratio between long and short intervals in the drum pattern and the number of counts per measure.

Half of the infants were familiarized with one of the two isochronous meter excerpts; the other infants were familiarized with one of the two non-isochronous meter excerpts. All stimuli were prepared as Quicktime movies accompanied by identical visual (non-rhythmic) portions of a documentary film (Attenborough, 1991).

3.2.1.3. Apparatus and Procedure

We tested infants by means of a familiarization-preference procedure. Infants sat on their parents' lap in a dimly lit testing room, with two monitors located approximately 140 cm in front and to the right and left of the infant. An observer recorded infant looking times by pressing one of two buttons on the computer, one button for looking towards a monitor, the other for looking away. To attract infants' attention to the appropriate monitor, a flashing red screen on that monitor preceded each trial. Infants were first presented with 2 min of the familiarization stimulus, consisting of four 30-s repetitions alternating between monitors. After four familiarization trials, test stimuli were presented six times each, with the structure-preserving and structure-disrupting test stimuli alternating between monitors. Each test trial was terminated when the infant looked away for 2 s or when 60 s had elapsed. The order of the first monitor in the familiarization phase, first monitor in test phase, and the first test stimulus presented (structure-preserving vs. structure-disrupting) was counterbalanced.

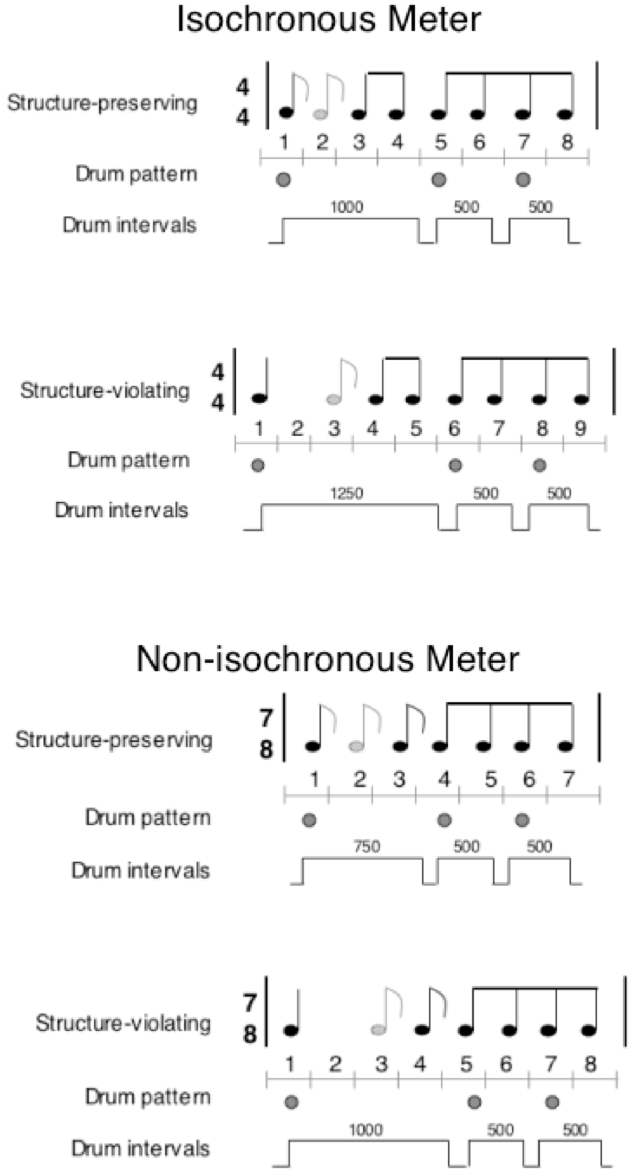


Figure 3.2. One measure of structure-preserving and structure-violating variations on isochronous and non-isochronous excerpts. A single extra note (in gray) is inserted into each measure for both types of variations. In structure-preserving variations, the note preceding the change is shortened to maintain the meter and the duration of intervals in the drum pattern. In structure-violating variations, no existing note

durations are modified, resulting in an increased long interval in the drum pattern and an extra count.

3.2.1. Results and Discussion

Infant looking time, which corresponds to listening time because of the contingency between looking and sound presentation, was accumulated over successive presentations of structure-preserving and structure-disrupting test stimuli. In the isochronous familiarization condition (Figure 3.3), looking (i.e., listening) times for the structure-disrupting variation (22.67 s) exceeded those for the structure-preserving variation (18.37 s) (paired two-tailed *t* test, $t = 3.50$, df = 25; $p < .01$), which is consistent with infants' typical preference for novel stimuli and replicates our prior findings (Chapter II). In the non-isochronous condition, however, looking times did not differ for structure-disrupting (24.54 s) and structure-preserving (24.3 s) variations (paired two-tailed *t* test, $t = 0.16$, df = 25; $p > .87$). This result contrasts with previous findings from 6-month-old listeners, who distinguished structure-disrupting from structure-preserving variations in the context of non-isochronous as well as isochronous meters (Chapter II). The present results are consistent with early developmental changes in speech and face perception. Infants' ability to differentiate foreign, non-isochronous rhythmic patterns declines by the end of the first year, but their sensitivity to comparable distinctions in culturally typical isochronous patterns remains unchanged.

If this developmental change arises from exposure to Western music during the first year of life, then exposure to music with foreign musical meters may prevent or reverse this apparent decline in sensitivity to foreign-meter variations. For example, after accumulating 5 hours of interactive experience (distributed over a 1-month period) with a native speaker of Mandarin Chinese, American 12-month-olds show sensitivity to Mandarin speech contrasts comparable to that of 6-month-olds, unlike 12-month-olds who receive no such exposure (Kuhl, Tsao, & Liu, 2003). In Experiment 2 we assessed 12-month-olds' ability to distinguish structure-disrupting from structure-preserving variations of foreign-meter tunes after brief, daily exposure to Balkan folk music.

3.3. Experiment 2

The goal of Experiment 2 was to assess whether brief, at-home exposure to non-isochronous meters in Balkan folk music could improve 12-month-olds' differentiation of structure-disrupting and structure-preserving variations in the foreign, non-isochronous context.

3.3.1. Method

3.3.1.1. Participants

Participants were 26 infants who were 11-12 months of age (*M* age = 350.6 days, *SD* = 16.4 days), 9 girls and 17 boys. All participants were healthy, full-term infants who were free of colds on the test day and had no family history of hearing impairment. Infants were living in a primarily monolingual English-language environment. Neither parents nor infants had prior exposure to Balkan music. An additional 10 infants were excluded from the final sample due to fussing (n = 9) or prior exposure to Balkan music (n = 1).

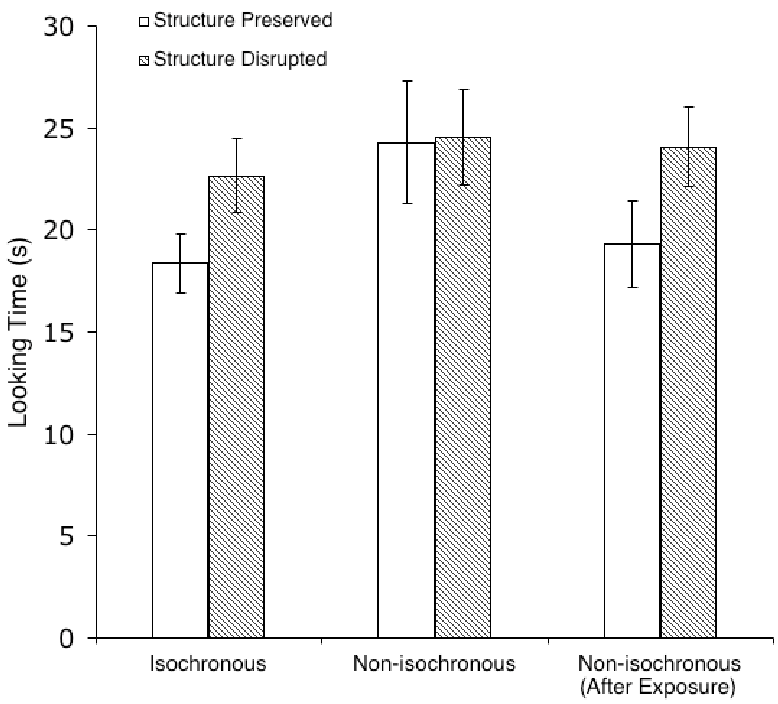


Figure 3.3. Infants' looking times (in seconds) during presentation of structure-violating and structure-preserving variations of isochronous (familiar) and non-isochronous (foreign) folk tunes in Experiments 1 and 2. Error bars represent standard errors.

3.3.1.2. Stimuli

A CD was prepared for at-home listening. Each CD was approximately 10 minutes in duration and contained five recordings of non-isochronous dance music from Macedonia, Bulgaria, or Bosnia (see Appendix for a detailed description of selections on the audio CD). The CD contained none of the tunes used during testing, so the metrical structure was the only feature common to melodies on the CD and those used during testing. Audio CDs, which were prepared by means of SoundEdit and iTunes, were burned onto CD with a Macintosh computer. Familiarization and test stimuli were identical to those used in the complex meter condition of Experiment 1.

3.3.1.3. Apparatus and Procedure

Two weeks prior to a laboratory visit, we mailed parents an audio CD, instructions, and a log sheet. Parents were instructed to play the CD at an audible level once every morning and afternoon when the infant was alert and contented. Moreover, they were asked to maintain the infants' regular routines, without drawing attention to the music. Parents kept track of music-listening sessions using the log sheet. After the two-week listening period, infants participated in a laboratory test session (non-isochronous condition). The procedure during test sessions was identical to Experiment 1.

3.3.2. Results and Discussion

Infants looked significantly longer during the structure-disrupting variation (24.07 s) than during the structure-preserving variation (19.31 s) (Figure 3.3) (paired two-tailed *t* test, $t = 2.297$, df = 1, 25; $p < .01$). A comparison of looking times in Experiment 2 and in the non-isochronous condition of Experiment 1 revealed a significant interaction between variation type (structure-disrupting vs. structure-preserving) and condition (exposure vs. no exposure) (mixed design analysis of variance, $F = 4.19$, df = 1,50; $p < .05$), indicating that 12-month-olds in Experiment 2 distinguished rhythmic variations in the foreign metrical context, unlike their peers in Experiment 1 who had no prior exposure to foreign-meter music. Moreover, preference scores (proportion of total looking time during the structure-disrupting variation) did not differ between the non-isochronous condition of Experiment 2 and the isochronous condition of Experiment 1 (two-tailed independent-samples *t* test, $t = 0.56$, df = 50; $p > .57$), suggesting that exposure led to native-like performance in the foreign-meter context. In short, two weeks of passive, at-home exposure facilitated infants' differentiation of rhythmic patterns in a foreign musical context.

3.4. Experiment 3

The goal of Experiment 3 was to determine whether comparable exposure to foreign music in adulthood could lead to comparable rapidity of learning. Adults were tested by means of a similarity judgment task. In this task, adults heard the same sequence of familiarization and test stimuli as did infants, but they responded by rating each variation on the basis of its similarity to the original familiarization stimulus. Adults were randomly assigned to an experimental or control group. Adults in the experimental group participated in two test sessions separated by 1 or 2 weeks, and they listened daily to an audio CD of Balkan folk music between the two test sessions. They could earn a modest monetary reward for accurately recognizing tunes from the CD in a recognition test following the second test session. Adults in the control group participated in two test sessions one or two weeks apart, with no exposure to non-isochronous music during the intervening period.

3.4.1. Method

3.4.1.1. Participants

Participants were 40 college students (25 women, 15 men, 18-35 years) whose musical training ranged from 0-15 years (*M* = 3.21). The mean duration of musical training was comparable across participants in the control condition (*M* = 3.38) and the experimental condition (*M* = 3.05). Most participants had grown up in North America, but some individuals had lived in Ireland (2 years), Sri Lanka (2 years), United Arab Emirates (1 year), Ukraine (10 years), Poland (3 years), and Pakistan (10 years). Individuals who had lived outside of North America were divided evenly between control and experimental groups. No participants reported prior exposure to Balkan music. An additional 6 participants were tested but not included in the final sample because they scored less than 75% on the recognition test.

3.4.1.2. Stimuli

Familiarization and test stimuli were identical to those used in Experiments 1 and 2. During the test phase, two additional test stimuli served as foils. One of these foils was an unaltered version of the primary melody from the familiarization stimulus (highly similar). The other was disrupted by the pseudo-random insertion of extra notes twice per measure (highly dissimilar). Foils were included to encourage participants to use the full range of the similarity judgment scale.

A brief set of practice trials preceded testing at the first session. For practice, we generated one familiarization stimulus and four test stimuli based on a children's tune (*Mary Had a Little Lamb*). Two structure-preserving test stimuli presented an identical rendition of the original, or contained extra notes that preserved the meter. Two structure-violating variations contained extra notes or pauses inserted several times per measure, resulting in salient disruptions of the metrical structure.

An audio CD for adult home listening was identical to the audio CD from Experiment 2, with the exception of one adult tune that replaced a children's tune (see Appendix). The audio CD was approximately 12 minutes in duration.

A recognition test was prepared for participants in the experimental condition. This test consisted of 20-s excerpts, five targets (taken from the audio CD) and 15 non-targets, 12 of which were drawn from the same Balkan artists. Thus, the voices, instruments, and style of most non-target excerpts were similar to target excerpts, precluding success on the recognition test without the requisite listening experience. The other 3 excerpts consisted of folk music and jazz and were included as obvious non-targets.

3.4.1.3. Apparatus and Procedure

Adults, who were tested alone or in pairs, listened to the stimuli over headphones at individual computer stations. PsyScope software presented stimuli and instructions and collected responses (Cohen, MacWhinney, Flatt, & Provost, 1993). Trials were presented in blocks consisting of one isochronous or non-isochronous 2-min familiarization stimulus followed by four test stimuli: two structure-preserving and two structure-disrupting variations of the familiarization stimulus, ordered randomly. Participants rated how well the four test stimuli matched the rhythmic structure of the familiarization stimulus (1, or very similar, to 6, or very dissimilar). Each block was repeated three times per session, resulting in three sets of judgments per stimulus per session. Block order was counterbalanced across participants.

Each participant was randomly assigned to the control or experimental group. Participants in the both groups were asked to return for a second session at the same time one (n = 22) or two (n = 18) weeks later. Participants in the experimental group were instructed to listen at home twice daily to the audio CD during the interim, in preparation for a subsequent recognition test. After completing the similarity judgment task at Session 2, participants in the experimental condition took the recognition test, presented by means of PsyScope (Cohen et al., 1993). They listened to a series of randomly ordered excerpts (targets and non-targets) and decided whether each one was or was not present on the audio CD. They could earn up to $10 for accurate recognition scores (percentage of hits and correct rejections).

Following the last session, participants completed a questionnaire assessing musical and cultural background, and, for participants in the experimental condition, the number of times they listened to the CD and the types of activities carried out while listening.

3.4.2. Results and Discussion

Higher dissimilarity ratings for structure-disrupting than for structure-preserving variations reflected accurate performance, so difference scores (structure-disrupting minus structure-preserving) provided a measure of accuracy on the similarity judgment task. Consistent with prior work (Chapter II), adults performed more accurately in the isochronous, Western conditions (1.79) than in the non-isochronous, Balkan conditions (-.20) (mixed design analysis of variance, $F = 74.61$, df = 1, 35; $p < .001$) (Figure 3.4). Note that accuracy scores in the non-isochronous conditions were generally below 0 (Figure 3.4), suggesting that adults did not just confuse the two variations, but actually performed the task incorrectly. In the non-isochronous condition, adults had a tendency to wrongly rate the structure-disrupting variations as more similar to the original rhythm than the structure-preserving variations. Because the structure-disrupting variation is consistent with Western meter, this pattern of performance likely reflects adults' tendency to assimilate the rhythms into a Western metrical framework.

Overall, adults in experimental and control groups performed similarly across the two sessions in the isochronous condition, but the experimental group performed more accurately in the non-isochronous condition of the second session, as shown by a three-way interaction between meter, session, and group (mixed design analysis of variance, $F = 4.50$, df = 1, 35; $p < .05$). At-home exposure thus generated some improvement, but accuracy remained at chance levels in the second session (two-tailed, one-sample t test, $t = 1.192$, df = 19; $p > .25$). In short, adults did not attain native-like levels of performance after exposure to foreign musical structures. This contrasts with 12-month-old infants, whose post-exposure performance in the foreign musical context was equivalent to their pre-exposure performance in the familiar musical context.

In principle, differences in the duration of exposure between sessions could account for greater learning in 12-month-olds than in adults. Although adults in Experiment 3 were instructed to listen to the CD twice daily, their self-reports indicated that, on average, they listened once daily ($M = 11.3$ min, $SD = 5.16$).

Duration of exposure did not, however, predict accuracy scores. We found no correlation between self-reported duration of daily listening and improvement ($r = -0.11$, df = 20, $p > .64$), nor did we find a difference between one- vs. two-week exposure groups (independent samples *t*-test, two-tailed, $t = 0.1$, df = 18, $p > .97$). Moreover, because adults were motivated by monetary reward for identifying tunes from the CD, they may have listened in an active, deliberate manner, which would have given them an advantage over infants, for whom the music played in the background during other activities. It is thus likely that differences between adults and 12-month-olds in post-exposure performance arose not from differences between the two learning contexts, but from age-related changes in the ability to learn foreign musical structures from perceptual experience.

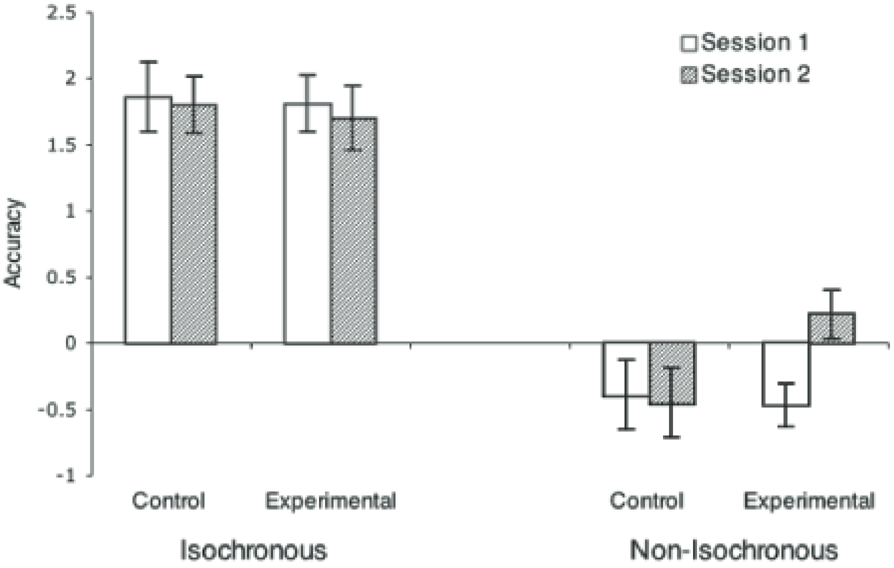


Figure 3.4. Mean accuracy scores (rating for structure-disrupting version minus rating for structure-preserving version) across sessions in the context of isochronous and non-isochronous meters. Only the accuracy of participants in the experimental group

improved from Session 1 to Session 2 in the complex meter condition. Error bars represent standard errors.

3.5. General Discussion

Taken together, Experiments 2 and 3 indicate that adults do not learn about foreign metrical structures as readily as do infants. Adults also have difficulty differentiating non-native speech sounds (Flege, Yeni-Komshian, & Liu, 1999; Takagi, 2002) and faces from unfamiliar racial groups (Ng & Lindsay, 1994; Goldstein & Chance, 1985) even after extensive experience with foreign speech and faces. Our findings may indicate that age-related constraints influence the nature and extent of learning from perceptual experience across multiple domains.

The results of Experiments 2 and 3 support two claims. First, passive exposure to the music of a particular culture can lead to the emergence of culture-specific responsiveness during infancy, presumably because infants learn rapidly from statistical regularities in input, whether that input is linguistic or musical (Kuhl, Tsao, Liu, Zhang, & De Boer, 2001Kuhl et al., 2003; Maye, Werker, & Gerken, 2002; Saffran, Aslin, & Newport, 1996). Second, perceptual experience may have a greater impact on musical rhythm perception during infancy than it does during adulthood. In principle, the adult-like responses of 12-month-olds in Experiment 1 could arise from adult-like representations. Our findings from Experiments 2 and 3 indicate, however, that adults' and infants' representations differ in fundamental ways. A lifetime of exposure to Western music may entail greater perceptual "commitment" to Western metrical structures, which results in foreign patterns being assimilated into stable and entrenched representations of familiar structures, one consequence of which is slower learning (Flege et al., 1999; Kuhl, Tsao, Liu, Zhang, & De Boer, 2001; Palmeri, Wong, & Gauthier, 2004). By contrast, infants' representations of musical meter may be considerably less robust and, consequently, more susceptible to modification. This interpretation is consistent with the notion that infant and child behaviors arise from representations that emerge gradually and grow in strength over the course of development as a result of increasing experience and brain development (Munakata, 2001).

We documented the emergence of culture-specific biases for the perception of musical rhythm in infancy, and we showed that relatively limited amounts of passive exposure could reverse these biases in infants but not in adults. In combination with prior work on speech and face perception, our results support the notion that domain-general perceptual tuning fosters the acquisition of knowledge about socially meaningful stimuli in the environment, such as sights and sounds associated with important persons or events. Simple learning mechanisms may underlie this perceptual tuning process, but our findings are consistent with the view that infants in particular have a heightened capacity for perceptual learning. Experience-related changes in the strength of representations may give rise to the appearance of “sensitive” periods during development, when learning is particularly rapid and behavior is modified easily. Our findings indicate that such processes are not specific to language or to faces, but extend to other domains such as music. Early learning about faces and voices is often viewed as evidence of their social and biological significance. Either music must be added to the list of socially and biologically significant stimuli, or the phenomenon of rapid perceptual attunement coupled with early flexibility must be more widespread than is currently believed.

CHAPTER IV.
SUMMARY

4.1. Main Findings

The goal of this dissertation was to investigate the perception of musical rhythm and meter in the first year after birth. Chapter 1 asked the basic question of whether 7-month-old infants can perceive the underlying metrical structure of simple rhythmic patterns, and, once perceived, whether they can detect associations between that inferred metrical framework and pitch structure. Chapters II and III examined the enculturation of adults' and infants' metrical representations, first by contrasting adults' culture-specific perceptual biases with 6-month-olds' culture-general rhythm perception (Chapter II), and then by examining when and how culture-specific knowledge of meter develops (Chapter III).

The findings support the conclusion that young infants perceive rhythm and meter, which may provide a critical starting point for early learning about musical organization. Chapter I reported that infants perceive the underlying meter of simple rhythms, and that they can learn contingencies between particular pitches and strong or weak positions in the meter. In principle, these abilities could provide infants with a foundation for learning about pitch and rhythmic organization in music. Chapter II reported further evidence of infants' sensitivity to meter. After familiarization to a folk tune, infants show a preference for a variation that disrupts the underlying meter over a variation containing a comparable change but no metrical disruption. This ability appears to depend on age and on cultural background. At 6 months, Western infants differentiate variations in both familiar (Western) contexts and in foreign (Balkan) metrical contexts. Their behavior parallels that of Bulgarian or Macedonian adults who are accustomed to Western and Balkan meters, but contrasts with Western adults who are accustomed to Western meters only. Chapter III reported that 12-month-old infants, unlike 6-month-olds, distinguish variations only in the Western context, presumably because of cumulative exposure to Western meters over the first year of postnatal life. Despite these changes, 12-month-olds readily learn to distinguish variations in the foreign context after brief exposure to Balkan music. Adults also seem to learn from experience, but their level of performance for foreign stimuli remains far below that for familiar stimuli, which implicates age-related changes in the strength and flexibility of culture-specific metrical representations. In short, this dissertation documented basic abilities for rhythm and meter perception early in development, but indicated that these abilities change rapidly in response to culture-specific listening experience.

There are common threads throughout the chapters of this dissertation. All studies reinforce the notion that general perceptual mechanisms, such as the ability to detect frequently occurring events or relations, enable infants to parse, remember, and learn about complex auditory input such as music and speech. Moreover, such mechanisms may be crucial for the acquisition of culture-specific knowledge. Recent work in the domain of speech perception emphasizes infants' capacity to learn complex structures from relatively limited perceptual experience (Kuhl, Tsao, & Liu, 2003; Maye, Werker, & Gerken, 2002). Instead of allowing individuals to register the world passively, perceptual experiences may actively tune representations and knowledge in a manner that is increasingly consistent with structure in the world (Kuhl, 2004; Kuhl, Tsao, Liu, Zhang, & De Boer, 2001; Palmeri, Wong, & Gauthier, 2004). Just as infants learn from the statistical frequency of speech sounds in their language input (Anderson, Morgan, & White, 2003; Maye, Werker, & Gerken, 2002), they may also learn from statistical information in music. Chapter I illustrated infants' sensitivity to the relative frequency of events and accents at periodic positions over time, and, conversely, to particular pitch events at those positions. In Chapters II and III, it is likely that infants detected the distribution of ratios between adjacent temporal intervals (such as a 2:1 or a 3:2 ratio), using this information to develop culture-specific categories or knowledge about which ratios are most typical in Western music. Thus, relatively simple statistical structures may provide a framework for more elaborate, culture-specific learning throughout later development.

4.2. Implications and Future Research

4.2.1. Is Rhythm Learned Before Other Aspects of Music?

Rhythm and meter may be particularly salient structures for infants, despite their inability to coordinate movements in synchrony with music or with others. This precedence of perception over production is consistent with infants' ability to perceive various aspects of pitch and melodic structure (Chang & Trehub, 1977a; Schellenberg & Trehub, 1996; Trehub, Thorpe, & Morrongiello, 1987; Trehub, Thorpe, & Trainor, 1990) well before those aspects can be reproduced vocally. Of particular note, for the present purposes, are striking differences in the acquisition of culture-specific patterns of pitch and rhythm processing. The present investigation revealed culture-specific biases in rhythm perception by 12 months of age, which contrasts markedly with proposed culture-specific responses to harmonic structure by 6-12 years of age (Cuddy & Badertscher, 1987; Krumhansl and Keil, 1982; Trainor & Trehub, 1994). Thus, culture-specific responses to rhythmic structure may emerge far earlier than culture-specific responses to musical pitch structure.

What could account for earlier learning of rhythmic structure? One possibility is greater prenatal exposure to rhythmic information. Due to the low-pass filtering characteristics of the abdominal wall, some auditory structures are audible in utero, such as the rhythms of maternal heartbeat as well as music and speech (Abrams, Gerhardt, Huang, Peters, & Langford, 2000; Lecanuet, 1996). Fetuses and newborns respond differentially to familiar over unfamiliar music (Kisilevsky et al., 2004) and speech (Moon, Cooper, & Fifer, 1993; Mehler, Jusczyk, Lambertz, Halsted, Bertoncini, & Amiel-Tison, 1988), which implies that rhythm learning may begin in the months prior to birth. A second possibility is that caregivers highlight rhythmic structures through their singing, bouncing, rocking, and communicative turn-taking (Papousek & Bornstein, 1992; Phillips-Silver & Trainor, 2004; Trainor, Clark, Huntley, & Adams, 1997). Caregivers may also enhance the rhythmic regularity of their infant-directed speech and singing in a didactic manner (Stern, Spieker, Barnett, & MacKain, 1983) or as a means of modulating infants' arousal (Trehub & Nakata, 2001). In addition to rhythm, many pitch structures (such as fundamental frequency) are also available prenatally (Lecanuet, 1996), and a substantial body of research indicates that pitch structure is essential for attracting infants' attention towards caregiver speech and singing (Fernald, 1992; Fernald & Kuhl, 1987). Thus, there are many unanswered questions about the relative salience of rhythm and pitch as well as the origins and consequence of any differential salience.

4.2.2. When Do Adult-like Representations of Rhythm and Meter Emerge?

The present work emphasizes the development of rhythm and meter perception during infancy, but it is likely that these abilities continue to develop throughout childhood. Developmental changes have been documented in children's spontaneous tapping and synchronous tapping to music (Drake, Jones, & Baruch, 2000) as well as their classification and discrimination of metrical versus nonmetrical patterns (Wilson, Wales, & Pattison, 1997). It is therefore likely that the perception of meter becomes increasingly adult-like over the course of development.

At 12 months, Western infants exhibited diminished sensitivity to variations of foreign, Balkan stimuli but not to Western stimuli (Experiment 1, Chapter III), which parallels the performance of Western adults (Experiment 1, Chapter II). It is important to emphasize that culture-specific responses to tasks in the present study do not translate to fully developed, adult-like perception of musical meter, especially by 12-month-old infants. This point is underscored by the different outcomes that resulted from at-home exposure to foreign music by 12-month-olds and adults. Recall that 12-month-olds performed at native-like levels in the foreign context but Western adults performed at chance levels in that context (Experiments 2 and 3, Chapter III). This divergent performance may stem from fundamentally different metrical representations in infants and adults. It is likely that infants have weaker representations of culture-specific metrical structures that are readily modified in response to foreign input. Adults, by contrast, may have much stronger biases for familiar structures at the onset of exposure to foreign music, which could interfere with their ability to achieve comparable performance levels in foreign contexts.

This interpretation is consistent with outcomes in the domain of speech perception. Although adults can learn to discriminate non-native speech sounds (Lively, Logan, & Pisoni, 1993; Logan, Lively, & Pisoni, 1991), their performance remains modest, typically well below native levels even after 1-6 weeks of intensive daily training (Callan et al., 2003; Lively et al., 1993; Miyawaki et al., 1975; Takagi, 2002). By contrast, very brief exposure during infancy can lead to native-like levels of discrimination performance in foreign speech contexts (Kuhl et al., 2003). Young learners may have advantages over mature learners because their knowledge and experience is less committed to native structures (Flege, Yeni-Komshian, & Liu, 1999; Kuhl et al., 2001). The findings of Chapter III imply similar origins for the observed age advantage in the context of musical rhythm learning.

In summary, enculturation processes have measurable effects on rhythm and meter perception during infancy, but undoubtedly these abilities continue to develop beyond infancy. Further research could investigate the consequences of exposure over a broad age range to determine when genuine adult-like representations emerge.

4.2.3. Parallels Between Rhythm in Music and Speech

Comparisons between developing music and speech perception skills were noted in various parts of this dissertation. The development of rhythm and meter perception parallels several important hallmarks in the development of speech perception. At young ages, infants can distinguish speech sounds from languages that they have never heard (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). Likewise, young infants can distinguish rhythmic variations in foreign metrical contexts (Chapter II). By 10-12 months of age, the ability to distinguish non-native speech sounds declines (Werker & Tees, 1984, 1999), as does the ability to distinguish rhythmic variations in a foreign context (Chapter III). Culture- and language-specific knowledge constrains both music and speech perception in adulthood. Adults fail to differentiate rhythmic variations in foreign but not familiar contexts (Chapter II). Likewise, language-specific knowledge can interfere with adults' discrimination of non-native speech sounds (Werker & Tees, 1984), and also with perception of stress patterns (Dupoux, Pallier, Sebastian, & Mehler, 1997) and consonant clusters (Dupoux Kakehi, Hirose, Pallier, & Mehler, 1999). An intriguing possibility is that the observed parallels extend beyond analogy, implicating underlying developmental processes that are common to both domains.

Although specialized, domain-specific processes may dominate adults' perception of music and speech, infants may approach both systems with similar perceptual skills and learning mechanisms (McMullen & Saffran, 2004). Comparable developmental trajectories in music and speech perception could arise, in part, from the salience of rhythm and timing in both speech and music. Linguists group the languages of the world into rhythmic classes, with stress-timed languages such as English and German characterized by regular inter-stress intervals, and syllable-timed languages such as French and Spanish characterized by regular inter-syllable intervals (Abercrombie, 1967). Recent work has identified an acoustic basis for these classifications; a measure of variability between adjacent vowel and consonant durations, called "durational contrast," can successfully group languages into appropriate rhythmic classes (Grabe & Low, 2002; Ramus, Nespor, & Mehler, 1999). These rhythmic properties of language are thought to form the basis of newborn preferences for native over non-native speech (Moon et al., 1993), because newborns only succeed in differentiating two non-native languages when they fall into contrasting rhythmic classes (Nazzi, Bertoncini, & Mehler, 1998). Older infants show increasingly language-specific responding, eventually exhibiting a preference for native speech even when the other speech sample comes from a language within the same rhythmic class (Bosch & Sebastian-Galles, 1997; Nazzi, Jusczyk, & Johnson, 2000). If durational contrast forms the basis of language categorization, it would rely on the basic ability to detect relations between adjacent durations--an ability that is essential for perceiving and categorizing musical rhythms according to culture-specific metrical structures.

Recent evidence indicates that links may exist between rhythmic structures in speech and music. When durational contrast, as described above (Grabe & Low, 2002; Ramus et al., 1999), is applied to musical notes, values obtained for languages match those obtained for music (Patel & Daniele, 2003). Specifically, culturally representative samples of French and English music differ in the same rhythmic manner as do samples of English and French speech. This raises the possibility that rhythmic structure in music and speech may drive similar developmental changes. Future work could investigate whether infant- and child-directed music exaggerates the rhythmic features that are consistent with native speech rhythms, and whether young infants prefer musical rhythms that match native speech rhythms.

4.3. Conclusion

The work presented in this dissertation attempted to address the little-studied question of how infants perceive rhythm and meter in music. Although there is evidence that infants can perceive and discriminate rhythmic patterns (Demany, McKenzie, & Vurpillot, 1977; Chang & Trehub, 1977b; Morrongiello, 1984; Pickens & Bahrick, 1997; Trehub & Thorpe, 1989), little effort has been directed at infants' ability to infer the underlying metrical structure of music. This question is essential for understanding the development of musical abilities, because meter enables some of the most fundamental musical behaviors in adults, such as synchronized singing and dancing and ensemble performance. This dissertation also examined the effects of musical enculturation, an endeavor that may become increasingly difficult to study with the increasing dissemination of Western and "world" music. By understanding how music listening experience shapes developing perceptual abilities, we may gain insight into human perceptual and cognitive constraints on musical systems across cultures and throughout history. Moreover, we may begin to understand how and why organisms modify their behavior and knowledge in accordance with significant structures in the environment.

APPENDIX

Audio Listening CD for Experiment 2 (Chapter Three)

1. *Ruchenitsa* (Macedonia). Duration 2:04. Meter 7/8 (2 + 2 + 3). Traditional instruments. From Good (2002), track 12.

2. *Hami Shahasar* (Bosnia). Duration 2:57. Meter 7/8 (3 + 2 + 2). Voice and traditional instruments. From Slavonian Traveling Band (1999), track 10.

3. *Zikino Kolo* (Macedonia). Duration 2:18. Meter 7/8 (2 + 2 + 3). Traditional instruments. From Zlatne Uste Balkan Brass Band (1993), track 12.

4. *Tale Ognenovski* (Macedonia). Duration 2:15. Meter 7/8 (3 + 2 + 2). Traditional instruments. From Good (2002), track 3.

5. *Daga* (Bulgaria). Duration 1:25. Meter 7/8 (3 + 2 + 2). Voice and orchestra. From Riorih (2003), track 5.

Audio Listening CD for Experiment 3 (Chapter Three)

1. *Ruchenitsa* (Macedonia). Duration 2:04. Meter 7/8 (2 + 2 + 3). Traditional instruments. From Good (2002), track 12.

2. *Hami Shahasar* (Bosnia). Duration 2:57. Meter 7/8 (3 + 2 + 2). Voice and traditional instruments. From Slavonian Traveling Band (1999), track 10.

3. *Zikino Kolo* (Macedonia). Duration 2:18. Meter 7/8 (2 + 2 + 3). Traditional instruments. From Zlatne Uste Balkan Brass Band (1993), track 12.

4. *Tale Ognenovski* (Macedonia). Duration 2:15. Meter 7/8 (3 + 2 + 2). Traditional instruments. From Good (2002), track 3.

Idi Da Go Sakash (Macedonia). Duration 2:23. Meter 7/8 (3 + 2 + 2). Voice and traditional instruments. From Good (2002), track 14.

REFERENCES

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

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.

Anderson, J.L., Morgan, J.L., & White, K.S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*, 155-182.

Attenborough, D. (Producer). (1991). *The trials of life: Courting* [Motion picture]. England: BBC Television.

Baruch, C., & Drake, C. (1997). Tempo discrimination in infants. *Infant Behavior and Development, 20*, 573-577.

Bergeson, T.R. (2002). *Perspectives on music and music listening in infancy.* Unpublished doctoral dissertation, University of Toronto.

Best, C.T., McRoberts, G.W., Sithole, N.M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345-360.

Bharucha, J.J., & Pryor, J.H. (1986). Disrupting the isochrony underlying rhythm: An asymmetry in discrimination. *Perception & Psychophysics, 40,* 137-141.

Brown, S. (2003). Biomusicology, and three biological paradoxes about music. *Bulletin of Psychology and the Arts, 4,* 15-17.

Callan, D.E., Tajima, K., Callan, A.M., Kubo, R., Masaki, S., & Akahane-Yamada, R. (2003). Learning-induced plasticity associated with improved identification performance after training of a difficult second-language contrast. *NeuroImage, 19*, 113-124.

Chang, H., & Trehub, S.E. (1977a). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology, 24*, 324-331.

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

Christiansen, M.H., Allen, J., & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes, 13*, 221-268.

Clarke, E.F. (1987). Levels of structure in the organization of musical time. *Contemporary Music Review, 2,* 211-239.

Clarke, E.F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music (2nd Edition*), pp.473-500. New York: Academic Press.

Clayton, M. (2000). *Time in Indian music.* New York: Oxford University 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.

Collier, G.L. & Wright, C.E. (1995). Temporal rescaling of simple and complex ratios in rhythmic tapping. *Journal of Experimental Psychology: Human Perception & Performance, 21,* 602-627.

Creel, S. C., & Newport, E. L. (2002). Tonal profiles of artificial scales: Implications for music learning. In Stevens, C., Burnham, D., McPherson, G., Schubert, E., & Renwick, J. (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, 2002.* Adelaide: Causal Productions.

Cuddy, L.L., & Badertscher, B. (1987). Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. *Perception & Psychophysics, 41,* 609-620.

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*, 145-171.

Curtin, S., Mintz, T.H., & Christiansen, M.H. (in press). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*.

Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua, 92*, 81-104.

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

Desain, P. & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception, 32,* 341-365.

Drake, C. (1998). Psychological processes involved in the temporal organization of complex auditory sequences: Universal and acquired processes. *Music perception, 16*, 11-26.

Drake, C., Jones, M.R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition, 77,* 251-288.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18*, 1-23.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1568-1579.

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A distressing "deafness" in French? *Journal of Memory and Language, 36*, 406-421.

Flege, J.E., Yeni-Komshian, G.H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78-104.

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H.Papousek, U. Jurgens, & M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches.* New York: Cambridge University Press.

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*, 279-293.

Eimas, P.D., Siqueland, E.R., Jusczyk, P.W., & Vigorito, J. (1971). Speech perception in infants. *Science, 171,* 303-306.

Essens, P., & Povel, D. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics, 37*, 1-7.

Foster, N.A., & Valentine, E.R. (2001). The effect of auditory stimulation on autobiographical recall in dementia. *Experimental Aging Research, 27,* 215-228.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149-180). New York: Academic Press.

Geisler, R. (1989). *The Bulgarian and Yugoslav collections*. Grass Valley, CA: The Village and Early Music Society.

Goldstein, A.G., & Chance, J.E. (1985). Effects of training on Japanese face recognition: Reduction of the other-race effect. *Bulletin of the Psychonomic Society, 23*, 211-214.

Good, A. (2002). *Dances of Macedonia and the Balkans* [CD]. Cleveland, OH: Folksounds Records.

Grabe, E., & Low, E.L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & M. Warner (Eds.), *Laboratory Phonology*. Berlin: Mouton de Gruyter, 515-546.

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*.

Horowitz, F.D., Paden, L., Bhana, K., & Self, P. (1972). An infant-control procedure for studying visual fixations. *Developmental Psychology, 7*, 90.

Huron, D., & Royal, M. (1996). What is melodic accent? Converging evidence from musical practice. *Music Perception, 13, 489-516*.

Järvinen, T., & Toiviainen, P. (2000). The effect of metre on the use of tones in jazz improvisation. *Musicae Scientiae, 4*, 55-74.

Jones, M.R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology, 9,* 193-228.

Jones, M.R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*, 313-319.

Jones, M.R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459-491.

Jusczyk, P.W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences, 3*, 323-328.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159-207.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*, B35-B42.

Kisilevsky, B.S., Hains, S.M.J., Jacquet, A.Y., Graniere-Deferre, C., & Lecanuet, J.P. (2004). Maturation of fetal responses to music. *Developmental Science, 7*, 550-559.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Krumhansl, C.L., & Keil, F.C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition, 10*, 243-251.

Kuhl, P.K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831-843.

Kuhl, P.K., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy:

Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100,* 9096-9100.

Kuhl, P.K., Tsao, F., Liu, H., Zhang, Y., & De Boer, B. (2001). Language/culture/mind/brain: Progress at the margins between disciplines. *Annals of the New York Academy of Sciences, 935*, 136-174.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255,* 606-608.

Large, E.W. (2000). Rhythm categorization in context. In C. Woods, G.B. Luck, R. Brochard, S.A. O'Neill, & J.A. Sloboda (Eds.), *Proceedings of the 6th International Conference on Music Perception and Cognition*. Keele, Staffordshire, UK: Department of Psychology. CD-ROM.

Large, E.W., & Jones, M.R. (1999). The dynamics of attending: How we track time-varying events. *Psychological Review, 106*, 119-159.

Large, E.W., & Kolen, J.F. (1994). Resonance and the perception of musical meter. *Connection Science, 6*, 177-208.

Lecanuet, J.-P. (1996). Prenatal auditory experience. In I. Deliège and J. Sloboda (Eds.), *Musical beginnings* (pp. 3-34). Oxford, England: Oxford University Press.

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

Lively, S.E., Logan, J.S., & Pisoni, D.B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America, 94,* 1242-1255.

London, J. (1995). Some examples of complex meters and their implications for models of metric perception. *Music Perception, 13,* 59-77.

Longuet-Higgins, H.C. & Lee, C.S. (1982). The perception of musical rhythms. *Perception, 11*, 115-128.

Lynch, M.P., Eilers, R.E., Oller, D.K., & Urbano, R.C. (1990). Innateness, experience, and music perception. *Psychological Science, 1,* 272-276.

Mattys, S.L., & Jusczyk, P.W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition, 78*, 91-121.

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.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J.J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18,* 331-340.

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

Morrongiello, B.A. (1984). Auditory temporal pattern perception in 6- and 12-month-old infants. *Developmental Psychology, 20*, 441-448.

Morrongiello, B.A., & Trehub, S.E. (1987). Age-related changes in auditory temporal perception. *Journal of Experimental Child Psychology, 44*, 413-426.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences, 7*, 309-315.

Myers, J., Jussczyk, P.W., Kemler-Nelson, D.G., Charles-Luce, J., Woodward, A.L., & Hirsh-Pasek, K. (1996). Infants' sensitivity to word boundaries in fluent speech. *Journal of Child Language, 23,* 1-30.

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.

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.

Newport, E.L., & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127-162.

Ng, W.J., & Lindsay, R.C.L. (1994). Cross-race facial recognition: Failure of the contact hypothesis. *Journal of Cross-Cultural Psychology, 25*, 217-232.

Palmer, C., Jungers, M.K., & Jusczyk, P.W. (2001). Episodic 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.

Palmer, C., & Pfordresher, P.Q. (2003). Incremental planning in sequence production. *Psychological Review, 110,* 683-712.

Palmeri, T.J., Wong, A.C-N., & Gauthier, I. (2004). Computational approaches to the development of perceptual expertise. *Trends in Cognitive Sciences, 8*, 378-386.

Papousek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In I. Deliege and J. Sloboda (Eds.), *Musical beginnings: Origins and development of musical competence*. New York: Oxford University Press.

Papousek, H., & Bornstein, M.H. (1992). Didactic interactions: Intuitive parental support of vocal and verbal development in human infants. In H.Papousek, U. Jurgens, & M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches.* New York: Cambridge University Press.

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

Pascalis, O., de Haan, M., & Nelson, C.A. (2002). Is face processing species-specific during the first year of life? *Science, 296,* 1321-1323.

Patel, A.D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 7,* 674-681.

Phillips-Silver, J., & Trainor, L.J. (2004). Music is a multisensory experience for infants and adults. Poster presented at the *5th International Multisensory Research Forum, Spain.*

Pickens, J., & Bahrick, L.E. (1997). Do infants perceive invariant tempo and rhythm in auditory-visual events? *Infant Behavior and Development, 20*, 349-357.

Povel, D. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3-18.

Povel, D.J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2,* 411-440.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 3*, 265-292.

Rice, T. (1994). *May it fill your soul: Experiencing Bulgarian music.* Chicago: University of Chicago Press.

Riorih (2003). Daga (“Rainbow”). On *38 songs for the little ones* [CD]. Sofia: Baiko Kanev.

Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes.* New York: Oxford University Press.

Saffran, J.R., Aslin, R.N., & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Sansavini, A., (1997). Neonatal perception of the rhythmical structure of speech: The role of stress patterns. *Early Development and Parenting, 6*, 3-13.

Schellenberg, E.G. (2002). Asymmetries in discrimination of musical intervals: Going out-of-tune is more noticeable than going in-tune. *Music Perception, 19*, 223-248.

Schellenberg, E.G., & Trehub, S.E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7*, 272-277.

Slavonian Traveling Band (1999). *Balkan mixed salad* [Slavonian Travelling Band] [CD]. New York: The Orchard.

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

Stern, D.M., Spieker, S., Barnett, R.K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language, 10,* 1-15.

Strogatz, S. (2003). *Sync: The emerging science of spontaneous order*. New York: Hyperion.

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America, 111*, 2887-2896.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, *39*, 706-716.

Thompson, W.F., Schellenberg, G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science, 12*, 248-251.

Toiviainen, P., & Snyder, J.S. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception, 21*, 43-80.

Trainor, L.J., Clark, E.D., Huntley, A., & Adams, B.A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development, 20*, 383-396.

Trainor, L.J., & Trehub, S.E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 394-402.

Trainor, L.J., & Trehub, S.E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception & Psychophysics, 56*, 125-132.

Trehub, S.E., Cohen, A.J., Thorpe, L.A., & Morrongiello, B.A. (1986). Development of the perception of musical relations: Semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 295-301.

Trehub, S.E., & Nakata, T. (2001). Emotion and music in infancy. *Musicae Scientiae, Special Issue*, 37-61.

Trehub, S.E., Schellenberg, E.G., & Kamenetsky, S.B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 965-975.

Trehub, S.E., Schneider, B.A., & Henderson, J.L. (1995). Gap detection in infants, children, and adults. *Journal of the Acoustical Society of America, 98*, 2532-2541.

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., Thorpe, L. A., & Morrongiello, B. A. (1985). Infants' perception of melodies: Changes in a single tone. . *Infant Behavior and Development, 8*, 213-223.

Trehub, S. E., Thorpe, L. A., & Trainor, L. J. (1990). Infants' perception of *good* and *bad* melodies. *Psychomusicology, 9*, 5-19.

Trehub, S.E., Trainor, L.J., & Unyk, A.M. (1993). Music and speech processing in the first year of life. *Advances in Child Development and Behavior, 24*, 1-35.

Werker, J.F., & Lalonde, C.E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24,* 672-683.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and*

Development, 7, 49-63.

Werker, J.F., & Tees, R.C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509-535

Wilson, S. J., Wales, R. J., Pattison, P. (1997). The representation of tonality and meter in children aged 7 and 9. *Journal of Experimental Child Psychology, 64,* 42-66.

Zlatne Uste Balkan Brass Band (1993). *No strings attached* [CD]. Cambridge, MA: Rounder Records.