# The Effect of Structural and Performance Factors in the Perception of Anacruses

Justin London
*Carleton College*

Tommi Himberg and Ian Cross
*University of Cambridge, Cambridge, United Kingdom*

**when a melody begins with an anacrusis,** (i.e., "pick up" notes), rhythm and meter are out of phase. Three experiments were conducted to investigate the interactions between structural (rhythm and pitch) and performance (articulation and tempo) factors on the perception of anacruses. The independent variables were rhythmic figure, initial melodic direction, initial melodic interval, implied harmony, articulation, and tempo. Participants tapped "every other beat" to melodies composed for each experiment; the phase-alignment of taps with the stimulus was the dependent measure of anacrustic vs. non-anacrustic perception. Experiment 1 found a strong main effect for rhythmic figure and an interaction between rhythmic figure and tempo. Experiment 2 showed that as tempo increased there was a systematic shift toward anacrustic perception of some melodies. Experiment 3 found that in a rhythmically impoverished context, pitch-based structural factors had only a weak effect on the perception of anacrusis.

*Received July 1, 2008, accepted July 23, 2009.*

**Key words:** meter, rhythm, accent, tempo, tapping

In Grosvenor Cooper and Leonard Meyer's (1960) landmark book *The Rhythmic Structure of Music*, the authors classified musical rhythms according to a small number of basic rhythmic archetypes which they described in terms poetic feet. More importantly, Cooper and Meyer showed how those archetypes may be combined to form larger, hierarchically integrated structures. In laying out their rhythmic taxonomy, Cooper and Meyer also systematically examined how each rhythmic archetype interacts with various metric frameworks. For example, they discuss "Trochee in Duple Meter" (pp. 38-40) and then "Trochee in Triple Meter" (pp. 40-43), and so on. A number of points emerge from both the structure of their argument and their many examples:

- On the lowest hierarchic levels, metric and rhythmic accents tend to coincide, save in cases of syncopation;
- Metric units and rhythmic groups are almost always the same length at the level of the measure and motive;
- Metric units and rhythmic groups may or may not be congruent.

These factors are interrelated. If a rhythmic figure begins with an accented element (in Cooper and Meyer's terms, a Trochee, which is Strong-Weak, or a Dactyl, which is Strong-Weak-Weak), then group and measure boundaries will be in phase. Conversely, if the rhythmic accent occurs at the end or middle of a group (i.e., an Iamb, which is Weak-Strong, an Anapest, which is Weak-Weak-Strong, or an Amphibrach, which is Weak-Strong-Weak), when a melody begins with such a group the result is a phrase that begins with an anacrusis. That is, these phrases start with one or more notes before the first relatively strong accent; these notes (single or multiple) are often referred to as "pick up" or "upbeat" notes. When a melody begins with an anacrusis, rhythmic grouping structure and meter are out of phase.

Figure 1 shows one of Cooper and Meyer's (1960) examples of rhythmic/metric noncongruence, from the third movement of Mozart's String Quartet, K. 387 (N.B., This is from a list given in Cooper and Meyer, 1960, p. 48; it is noteworthy that all 11 of their examples of "iambs in duple and triple meter" in this list involve anacruses). The initial rhythmic group (dotted eighth-sixteenth-quarter note-quarter rest) is marked by the relatively long inter-onset interval (IOI) between its final note and the first note of the next group, a varied repeat of the first group. As a result of this gap, the group boundary is clear even before the parallel rhythmic structure emerges in the next measure. According to Cooper and Meyer's analysis, the dotted eighth-sixteenth note pair is weak relative to the following quarter note. Thus on the quarter-note level, the figure is iambic (weak-strong). This rhythmic group also establishes a sense of tactus on the quarter note level, given the Andante tempo.

*Music Perception* volume 27, issue 2, pp. 103–120, issn 0730-7829, electronic issn 1533-8312  DOI:10.1525/mp.2009.27.2.103


FIGURE 1. Example of an anacrustic melody, after Cooper and Meyer (1960).

When the figure repeats, the three beat metric cycle becomes immediately apparent. Note how the non-congruence between group and metric boundaries created by the initial anacrusis tends to persist in the remainder of the musical phrase (and indeed, may be characteristic of the grouping-meter relationship for an entire piece).

While Cooper and Meyer (1960) make a broad distinction between metrical accent versus a rhythmic "stress" (pp. 7-8), more recent theoretical treatments of accent are more nuanced. Lerdahl and Jackendoff distinguish phenomenal accents, structural accents, and metrical accents (Lerdahl & Jackendoff, 1983, p. 17). Accent also has been shown to be dependent on cultural factors (Toiviainen & Eerola, 2003). Different musical parameters can give rise to different kinds of accent, but in general, phenomenal accents are the product of the physical features of the musical stimulus (relative loudness, pitch height, etc.), while structural accents and metrical accents respectively involve syntactic judgements and metrical interpretations—that is to say, enculturated responses to standard structural configurations in a familiar musical style (for a further discussion of metrical accent see London, 2001, 2004).

Tonal motion—and most especially an implicit or explicit V to I harmonic motion—is thought in many music-theoretic accounts to be the most salient factor in determining rhythmic and/or metrical accent, and hence anacrustic structure (this is Yeston's, 1976, "pitch-to-rhythm" orientation of structure; see also Cone, 1968, and Rothstein, 1989). In an unaccompanied melody, this involves the harmonic implications of particular scale degrees, most especially $\hat{5} \rightarrow \hat{1}$ and $\hat{7} \rightarrow \hat{1}$, as well as figures that involve $\hat{5} \rightarrow \hat{7} \rightarrow \hat{1}$, $\hat{5} \rightarrow \hat{4} \rightarrow \hat{3}$, and so on (that is, scale degrees that can be regarded as members of a V or V7 chord followed by scale degrees that can be regarded as members of a I chord). Without a clear tonal framework that would allow for tones to be heard in a scalar context, certain intervals (ascending perfect 4ths or minor 2nds, descending perfect 5ths), as well as melodic direction (ascending figures in general) are presumed to bias listeners toward anacrustic perceptions.

Melodic contour or "melodic accent" has been the focus of several studies of melodic perception, most notably Huron and Royal (1996). They noted that in the music-theoretic literature there were "at least seven different conceptions of melodic accent . . . (1) treble accent [higher pitches more accented than lower], (2) bass accent [reverse of treble accent], (3) registral extreme accent [highest or lowest note in a passage], (4) interval size accent, (5) interval ascent accent, (6) interval descent accent, and (7) contour pivot accent" (Huron & Royal, 1996, p. 491). They explicitly do not consider tonality-related melodic accents (p. 491). After reviewing eight different conceptions of melodic accent, and making their own statistical study of accent in various musical corpuses, Huron and Royal concluded that "melodic accent may be a relatively weak factor in rhythmic perception and musical organization," though they also noted that "evidence of melodic accent appears to be most easily detected in unaccompanied isochronous solo passages" (p. 509). They further remarked that ". . . this study call[s] into question the notion that large intervals evoke melodic accents (p. 511).

Finally, the effect of durational structure on the perception of accent is well known in both the music-theoretic and psychological literature. It has long been observed that accent accrues to the first and/or (especially) last element in a rhythmic series and to relatively long durations or inter-stimulus onset intervals (Handel, 1998; Povel & Essens,1985; for a summary, see Krumhansl, 2000). Likewise Lerdahl and Jackendoff (1983), Benjamin (1984), Kramer (1988), Lee (1991), Drake (1993), and Hasty (1997) note the correlation between relative duration and rhythmic accent.

Returning to Figure 1, one can see how it involves a constellation of these accentual/anacrustic factors:

- It begins with several (two) shorter notes moving to a longer note;
- It involves a large, ascending interval;
- It implies a V-I harmonic motion in the key of C-major.

London (2004) makes two relevant observations on the perception of anacruses. The first is his notion of metric "malleability," a species of metric ambiguity. Metric malleability refers to the fact that many times a melodic sequence or durational pattern may afford a number of structural interpretations in terms of grouping, meter, and group-meter interactions, though in actual performance expressive variations may disambiguate its grouping and/or metric structure (London, 2004, pp. 48-50; see also Longuet-Higgins & Lee, 1984). At times this malleability involves whether or not a melody is heard as beginning with an anacrusis; Sloboda (1983) studied anacrustic and non-anacrustic performances of the "same" malleable melody. London's second observation is that durational patterns may behave differently (that is, are heard differently) in different temporal ranges. London characterized these temporal ranges in terms of "tempo-metrical types," as tempo affects the relationship between rhythmic elements that are heard as subdivisions of the beat, on the beat level, and on longer levels (London, 2004, pp. 76-79). When the global tempo of a melody is significantly increased, for example, the absolute values of various durations may shift to such an extent such that elements that were heard as articulating beats are now heard as subdivisions of the beat level. This framed the question for our research: is there a link between tempo (i.e., tempo-metrical type) and the listener's metrical construal of various melodic patterns?

The current study examined the effect of structural and performance factors on the perception of anacrusis—that is, on the accentual organization of the initial tones of an unaccompanied melody. By *structural factors* we mean durational pattern, melodic contour in terms of both interval direction and size, and the harmonies associated with particular scale degrees. By *performance factors* we mean overall tempo and between-group articulation (i.e., offset timing). Structural factors remain invariant under various performance transformations—thus, while one may sing "Happy Birthday" fast or slow, legato or staccato, certain intervallic and durational relationships must remain intact if these performances are all to count as instances of "Happy Birthday." Conversely, performance factors may change, and our study addresses those factors that are under a musician's conscious control, reflecting overt decisions he or she makes when approaching a performance.

## Experiments

In our experiments we focused on the opening three tones of each melody, as they establish the relation between grouping structure and meter. These tones were not studied in isolation, or in the context of a cyclically repeating figure (e.g., Drake, 1993; Povel & Essens, 1985), but as the initiation of a musically typical four-bar phrase. We thus considered accent in an explicitly musical context. Our basic contrast/primary variable was between melodies that begin with a Short-Short-Long (SSL) rhythm, in a 1:1:2 durational ratio, versus melodies which being with a Long-Short-Short (LSS) rhythm (again, a 2:1:1 ratio). Our other primary variable was that of tempo, to test London's conjectures regarding tempo-metrical types.

Given the correlations between relative duration and accent noted above, SSL melodies were presumably anacrustic, and LSS melodies presumably downbeat. We also hypothesized that melodies beginning with an initial ascent, melodic skip, V-I harmonic implication, and beginning on scale degree $\hat{5}$ would tend toward anacrustic perception, while melodies beginning with repeated tones, sustained tonic harmony, and beginning on scale degree $\hat{3}$ or $\hat{1}$ would tend toward non-anacrustic perception. Melodies that involve a mixture of these factors would tend to be metrically malleable, and hence anacrustically ambiguous. Figure 2 gives an example of a metrically malleable melody: while the SSL figure and melodic ascent are anacrustic factors, the harmonic stasis and start on scale degree $\hat{3}$ are non-anacrustic.

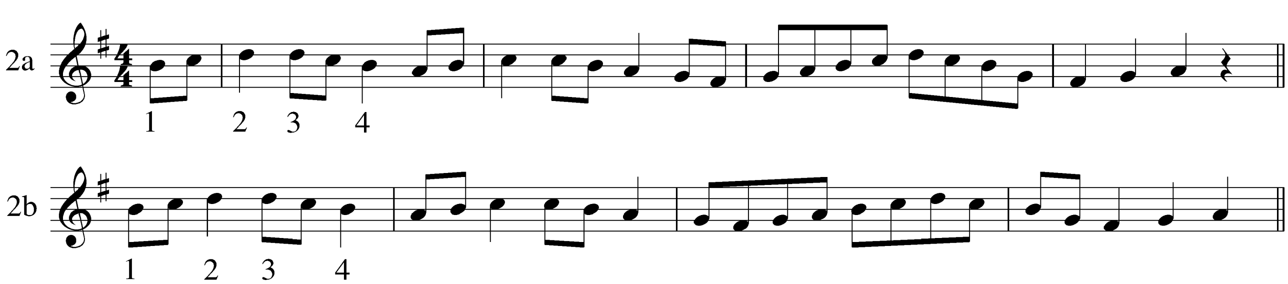


FIGURE 2. Upbeat (anacrustic) vs. downbeat interpretations of the same stimulus melody.

As indicated by the correspondence between numbers below the staff and the barline placement, in Figure 2a the first two tones are unaccented (hence anacrustic), while in 2b the "same" two tone pair is accented, and hence is heard commencing on the downbeat. Note that the melody in Figure 2 is ambiguous only in regard to downbeat location (i.e., anacrustic vs. non-anacrustic); the beat locations (relative to the eighth note subdivision), beat period, and measure period are unambiguous.

As we wished to study how the opening tones of a melody may affect metrical interpretation, over the course of these three experiments we systematically probed the following structural variables: *Rhythm* (SSL vs. LSS), *Melodic Direction* (Ascending vs. Descending vs. Repeated Tones), *Melodic Interval* ("Skip" [greater than a minor third] vs. "Step" [less than or equal to a minor third]), and *Implicit Harmony* (Initial V-I motion versus I-I prolongation). The performance factor of *tempo* was varied in Experiments 1 and 2, while *articulation* (staccato vs. legato) was investigated in Experiment 2. Moreover, stimuli were constructed so that their continuations were metrically neutral, i.e., that they would not introduce any reverse bias into the participants' initial metric construals of the opening figure. Specifically, melodies were composed to avoid sequential repetition or other patterning in their opening measures that might cue a particular rhythmic interpretation, and to then continue with largely stepwise motion. We also took care to construct the opening measures of all our melodies so that any skips between consecutive "S" elements would be harmonically sensible (i.e., members of either a I or V chord), and that all stepwise motions would be regarded as passing tones between stable chord members (again, between tones of a I or V chord). In addition, on the half note level, melodies were designed to project a stable harmonic structure in either anacrustic or non-anacrustic interpretations. In all three experiments all melodies were in G-major, so that tonality induction was not a factor over the course of the experiment. A complete set of the melodies used in the experiments is given in Appendices 1-3.

The task in all three experiments involved participants tapping "every other beat" while they listened to the melodic stimuli; a brief tutorial that demonstrated the task was presented prior to testing for each participant. We presumed that in this response mode participants would: (a) tap on the beat (i.e., they would not spontaneously engage in anti-phase tapping; see Longuet-Higgins & Lee, 1982; Repp, Iverson, & Patel 2008); and (b) align their taps with what they perceived (consciously or subconsciously) as the metrically accented tones in the pattern. Participants also were told to listen for 3-5 beats before they began tapping. This allowed for an observation of participants' perception of an anacrusis, as one can project the alignment of their tapping pattern in a particular trial backward to determine if they heard the sequence beginning with an anacrusis or not. We arrived at this response modality after using other forms of data collection in pilot studies, most often by having participants simply self-report (by means of a tally sheet) whether a given stimulus began with or without an anacrusis. However, we noted that participants often would rehearse and at times reconstrue a given stimulus after its presentation, switching between an anacrustic versus a non-anacrustic construal of the stimulus. This also introduced a memory component into the experiment. The "tap every other beat" response allowed for a natural and reasonably spontaneous response that avoided excessive ratiocination. Nonetheless, to insure that tapping responses would be indicative of metric perception, posthoc correlations were run between tap and tally responses for two sets of data: (a) the stimuli used in the second experiment, for which we had both tap and tally responses from all participants; and (b) a balanced set of LSS and SLL melodies that were used in both the first experiment (tap data) and in a separate pilot study (tally data). The correlations between tap-tally responses in these sets were both quite strong, $r(34) = .84$, $p < .001$ and $r(22) = .91$, $p < .001$, respectively. We thus feel reasonably secure that our "tapping every other beat" response represents a fairly transparent measure of anacrustic perception.

As the stimuli were presented via a MIDI sequencing program with deadpan timing, we indexed each melody relative to the nominal meter in the manner given in Figure 2b. It did not matter if a sequence involved LSS or SSL patterning, as this allowed us to refer to the first beat of every sequence as beat 1. All stimuli were unambiguously in 4/4 meter, so that whether anacrustic or not, all involved a 4 beat metric cycle. The experimental task, as noted above, was for participants to tap on "every other beat" after waiting 3-5 beats to gauge the metrical structure of each trial melody. This meant that their responses could be characterized as tapping on either the "odd" or "even" beats—the former indicating a non-anacrustic perception, the latter anacrustic, relative to the default indexing of beats noted above. Using a MATLAB script, every tap was associated with a beat in the music. This was done by dividing the timeline of the trial into "bins" that were one beat wide, centered on the exact time of the beat in the stimulus. So, for instance any tap occurring after the halfway point between beats 1 and 2 (see Figure 2), but before the halfway point between beats 2 and 3, would be associated with beat number 2.

The participant response for each trial was coded as either 0 (non-anacrustic) or 1 (anacrustic). A perfect response involved taps only on odd or even beats. In cases where there were occasional extra taps (accounting for less than 1/3 of the total taps), the trial was still coded as 0 or 1. In cases where the participant tapped on most or all of the beats, the data from that trial were rejected. These scores could then be averaged across all trials and all participants for a given stimulus. An average score of 0 indicated all participants tapped/heard the stimulus as non-anacrustic, while conversely an average score of 1 indicated uniform anacrustic perception; most scores, of course, fell between these extremes. In this way, upbeatness (i.e., tendency to be heard as anacrustic) could be measured for each stimulus. Similarly, scores for all responses in a given stimulus category were averaged, and those grand averages were then compared. This then permitted the relative contribution of various factors to be analyzed.

Experiments 1 and 3 were performed at Carleton College in Northfield, Minnesota. Experiment 2 and a pilot experiment were performed at the Centre for Music and Science at the University of Cambridge, UK.

## Experiment 1

### *Method*

#### PARTICIPANTS

Eighteen participants were involved in Experiment 1; none were excluded. Participants were recruited from the Carleton College community, including students, faculty, and staff. There were 10 men and 8 women; 16 were right-handed and 2 were left-handed (handedness was not deemed relevant for this task; participants were free to use whichever hand they wished for the tapping task, and all used their dominant hand). Their ages ranged from 17 to 63 years; average age was 29 years ($SD = 14.8$). Musical background was assessed in terms of three categories: less than 5 years of private study on a musical instrument or voice (6 participants), 5-10 years (4 participants), and more than 10 years (8 participants); the average years of musical study was 8.2 years ($SD = 6.4$).

#### STIMULI AND DESIGN

A $2 \times 2 \times 2 \times 2 \times 3$ factorial design was used: rhythmic figure (SSL vs. LSS), melodic direction (ascending vs. descending), melodic interval size (skip vs. step), implied harmony (V-I vs. I-I), and tempo (80 bpm/750 ms IOI, 105 bpm/570 ms IOI, and 140 bpm/428 ms IOI); the 16 basic sequences are given in Appendix 1. The 16 melodies at the 3 tempi yielded 48 test sequences. These were referred to as *melodic* stimuli. In addition, a set of *isotonic* stimuli (i.e., sequences that merely repeated the same note, akin to patterns played on a drum) also was constructed. This resulted in a simpler $2 \times 3 \times 3$ factorial design: rhythmic figure (SSL vs. LSS), scale degree (on scale degrees $\hat{1}$, $\hat{3}$, or $\hat{5}$), and tempo (same tempi as above), which yielded 18 additional test sequences. Hence, 66 sequences were employed in all.

The experiment used a within-subjects design; trials were conducted with individual participant sessions and sequences were presented in two different quasi-random orders. Stimuli were first randomised, and order of presentation manually adjusted to avoid consecutive presentations of either the same stimulus melody at different tempi or of runs longer than four stimuli at the same tempo. The two different orders of presentation were produced by splitting the 66 stimuli into two blocks A and B, each of 33 stimuli; half the participants heard the stimuli in order AB and half in order BA.

#### APPARATUS AND PROCEDURE

Experiment melodies were composed using the Sibelius 4.0 music notation program and then transferred to the Digital Performer sequencing program, running on a Macintosh G5 computer (OSX version 10.3.9). Digital Performer was used for both stimulus presentation and data collection (version 5.0 for experiments 1 and 3; version 4.1 for Experiment 2). In Experiments 1 and 3 the computer was linked via a M-Audio 2x2 MIDI interface to a Kurzweil K2600 Synthesizer and a Roland SPD6 drum pad. The "Concert Grand Piano" tone setting was used for the stimuli. Stimulus melodies were presented with deadpan timing and equalized loudness levels. Participants heard melodies over Sennheiser HD 280 Pro headphones adjusted to a comfortable listening level. The start of each stimulus was cued by the experimenter (JL) with a random interval of between 4 and 7 s between the presentation of each stimulus. The entire procedure took no more than 25 min.

Participants were told that they would hear 66 short (4-bar) melodies, and reminded that they should tap on every other beat. They were instructed that they should listen for 3-5 beats, and then start tapping. Some participants explicitly asked if they should tap "on 1 and 3, or on 2 and 4"; they were told to tap on 1 and 3, but also to avoid ratiocination or explicit counting procedures as much as possible.

## Results

Two separate analyses were conducted, one exploring melodic (ascending and descending) sequences and one exploring isotonic (repeated tone) sequences. Results for all three experiments are given in Table 1.

### *Melodic Sequences*

Preliminary analysis indicated no effect of either training, $F(2, 12) = 0.99$, $p > .10$, or of order, $F(1, 12) = 0.24$, $p > .10$, so results for both orders were collapsed and a repeated measures ANOVA was conducted on the data for ascending and descending sequences with rhythmic figure (2 levels), direction of initial interval (2 levels), initial harmony (2 levels), initial interval size (2 levels) and tempo (3 levels) as independent variables.

As was expected, there was a strong main effect for rhythmic figure (SSL vs. LSS), as the average upbeatness rating for SSL sequences was .79 while the upbeatness rating for LSS sequences was .12, $F(1, 17) = 83.70$, $p < .001$ (Greenhouse-Geisser correction applied). There was also a small, but significant main effect for melodic direction, $F(1, 17) = 4.98$, $p = .04$ (Greenhouse-Geisser correction applied), in that ascending produced slightly more upbeat responses than descending, with a difference from .47 for ascending melodies versus .44 for descending.

There was a significant two-way interaction between rhythmic figure and tempo, $F(1.60, 27.15) = 3.87$, $p = .04$, $\eta^2 = .23$ (Greenhouse-Geisser correction applied). Though the change is slight in both rhythmic conditions, as tempo increases the LSS patterns yield more downbeat responses while the SSL patterns educe more upbeat responses (see Figure 3).

Finally, there was a statistically significant four-way interaction of rhythmic figure (SSL vs. LSS) × melodic direction (Ascending vs. Descending) × implied harmony (V-I vs. I-I) × Interval Size (Skip vs. Step), $F(1, 17) = 5.94$, $p = .03$, $\eta^2 = .26$ (Greenhouse-Geisser correction applied). This appears to arise because of a slight increase in the extent to which a V-I ascent by a small interval in the context of an LSS rhythm induces an upbeat response compared with a descent. There was no clear effect of interval size, implied harmony, or melodic direction evident in the context of a SSL rhythm.

No other main effects or interactions were significant.

TABLE 1. Summary of ANOVA Results for Experiments 1, 2, and 3.

| | **Experiment 1** | | | | **Experiment 2** | | **Experiment 3** | |
|---|---|---|---|---|---|---|---|---|
| | **Melodic** | | **Isotonic** | | | | | |
| **Factors** | Levels | *F* | Levels | *F* | Levels | *F* | Levels | *F* |
| ***Structural factors*** | | | | | | | | |
| Melodic direction (asc., desc., none) | 2 | 4.98* | — | — | 3 | 7.44* | 2 | 5.97* |
| Implied harmony (V-I, I-I) | 2 | *n.s.* | 2 | 4.42* | | | 2 | 9.51** |
| Interval size (skip > m3d, step ≤ m3d) | 2 | *n.s.* | — | — | | | 2 | *n.s.* |
| Rhythmic figure (LSS, SSL) | 2 | 83.70*** | 2 | 10.99** | — | — | — | — |
| ***Performance factors*** | | | | | | | | |
| Tempo | 3 | *n.s.* | 3 | *n.s.* | 6 | 6.47** | 2 | *n.s.* |
| Articulation (staccato, legato) | 2 | — | | — | 2 | *n.s.* | | — |
| ***Participant factors*** | | | | | | | | |
| Training (< 5 yrs, 5-10 yrs, > 10 yrs) | 3 | *n.s.* | | *n.s.* | | *n.s.* | | *n.s.* |
| ***Significant interactions*** | | | | | | | | |
| Rhythmic figure × Tempo | | 3.87* | | 8.21** | | | | |
| Rhythm × Mel. dir. × Scale deg. × Interval size | | 5.94* | | | | | | |
| Melodic direction × Interval size | | | | | | | | 16.36** |
| Impl. harmony × Int. size × Training | | | | | | | | 15.02*** |

*Note*. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$; *n.s.* indicates a main effect that is not statistically significant.

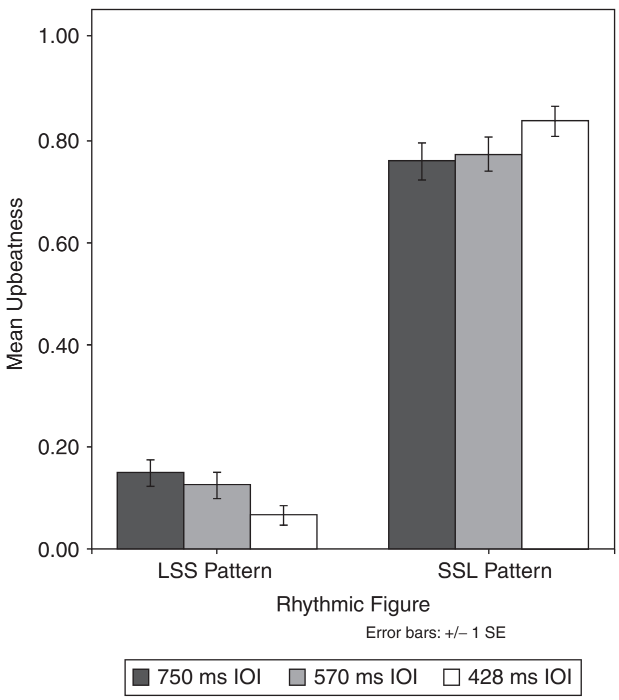


**FIGURE 3.** Interaction between rhythmic figure and tempo in ascending and descending contexts, Experiment 1.

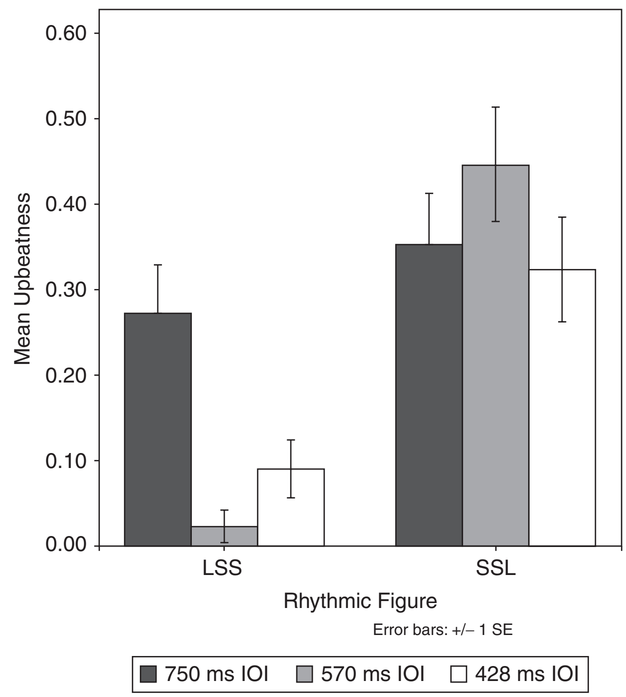


**FIGURE 4.** Interaction between tempo and rhythmic figure, isotonic context (Tempo 1 = 80 BPM; Tempo 2 = 105 BPM; Tempo 3 = 140 BPM), Experiment 1.

### *Isotonic Sequences*

There was no main effect of order, $F(1, 12) = 3.07$, $p > .1$, nor was there any effect of training, $F(2, 12) = 2.28$, $p > .1$, so results for both orders were collapsed and a repeated measures ANOVA was conducted on the data for isotonic sequences.

Again, there was a significant main effect for rhythmic figure: LSS had a mean upbeatness rating of .13, while SSL had a mean upbeatness rating of .38, $F(1, 17) = 11.00$, $p = .004$ (Greenhouse-Geisser correction applied). While the upbeatness rating of the LSS rhythm was commensurate with that found in the ascending and descending sequences (.12 vs. .13), the upbeatness ratings for the SSL sequences were considerably lower in the isotonic context (.38 vs. .79); see further discussion below.

Surprisingly, there was an effect of implied harmony, or in the case of these isotonic sequences, scale degree, $F(1.40, 23.75) = 4.42$, $p = .04$, $\eta^2 = .21$ (Greenhouse-Geisser correction applied), as it was presumed that these isotonic patterns would be heard much like unpitched drum patters or metronome clicks. Posthoc contrasts showed that sequences on scale degree $\hat{1}$ were marginally significantly less upbeat than the scaled degree $\hat{5}$ (upbeatness index of .21 vs. .30; $p = .06$, $\eta^2 = .20$, while those on scale degree $\hat{3}$ were significantly less upbeat than those on $\hat{5}$ (upbeatness index of .21 vs. .30; $p = .03$, $\eta^2 = .24$). There was no significant difference between sequences on scale degrees $\hat{1}$ and $\hat{3}$.

Again, there was also a statistically significant two-way interaction between rhythmic figure and tempo, a more pronounced effect than the one observed for the ascending and descending sequences, $F(1.74, 29.61) = 8.21$, $p = .002$ (Greenhouse-Geisser correction applied). In the isotonic context the slowest tempo (80 bpm) had the least effect on upbeatness in both LSS and SSL patterns, while at the other two tempi the contrast between the two rhythmic conditions was far more pronounced, as LSS produced consistently more downbeat responses and SSL more upbeat responses (see Figure 4). A series of paired $t$ tests confirmed that at the slowest tempo (750 ms), there was no significant difference in perceived upbeatness for the two rhythmic conditions, $t(17) = -0.90$, $p > .3$, whereas at the two other tempi, perceptions of upbeatness were significantly different for the two rhythmic conditions; for 105 bpm, $t(17) = -4.53$, $p < .001$, and for 140bpm, $t(17) = -2.94$, $p < .01$.

## Discussion

Experiment 1 confirmed, unsurprisingly, that durational structure is a primary determinant of metric accent, and hence whether a figure would be heard as anacrustic or non-anacrustic. But this does not mean that LSS patterns are non-anacrustic, while SSL patterns are anacrustic; the results from the trials with isotonic stimuli showed that without any melodic structure, both SSL and LSS patterns tend to be heard as non-anacrustic, as indicated by their upbeatness grand averages of .38 and .13, respectively. To put it another way, the addition of any melodic structure—whether a skip or step, tonic or dominant harmony, etc.—increases the upbeatness of both LSS and SSL patterns. This effect was quite dramatic in the case of SSL melodies, a jump from .38 to .79, indicating a fairly clear shift from downbeat to upbeat interpretation.

Tempo affected SSL and LSS melodies in different ways, as evidenced by their interaction. We believe this is related to the absolute duration of the elements in these figures at different tempos. As tempo increased in SSL contexts, the initial "S" elements get shorter (decreasing here from 375 ms to 285 ms to 214 ms); at the faster tempos the "S" elements are less likely to be regarded as beats and are more likely to be heard as beat subdivisions. When the "L" element of the SSL figure finally arrives, it effects a closure to the rhythmic figure (as per Narmour's characterization of a cumulative rhythm—Narmour, 1990, p. 105). At the same time, the pair of "S" elements defines the probable beat level IOI, and the following "L" confirms it. Thus at faster tempos the determination of beat and accent for SSL figures seems to follow the algorithms described by Longuet-Higgins and Lee (1984) and Desain and Honing (1999). Conversely, at slower tempos, LSS were less downbeat; presumably this is because at slower tempi the relatively long (750 ms) IOI that corresponds to the initial "L" element is less salient as a beat (Parncutt, 1994).

In sum, the first experiment showed that SSL patterns were more sensitive to tempo changes than LSS in terms of their accentual and anacrustic structure. While other structural factors did have an impact on upbeatness, the relative contribution of each was unclear, no doubt due to the large number of factors involved in this initial study. Therefore, additional experiments were designed to test the effect of tempo and structural factors, respectively, using simpler experimental designs.

## Experiment 2

### *Method*

#### PARTICIPANTS

Twenty-one participants were recruited for Experiment 2 from the University of Cambridge community. One participant's data were excluded due to inability to perform the experimental task. Of the remaining 20 participants, 9 were male and 11 female; 18 were right-handed and 2 left-handed. Their ages ranged from 21 to 66 years, with an average age of 33.3 years ($SD = 12.2$). Participants were categorized according to musical background as in Experiment 1: less than 5 years formal training (2 participants); 5-10 years (4 participants); and more than 10 years (14 participants). The average number of years of study was 13.0 years ($SD = 5.6$)

#### STIMULI AND DESIGN

Based on the results of Experiment 1, Experiment 2 was designed to give a more fine-grained assessment of the effect of tempo and articulation on anacrustic perception. To that end, a set of SSL melodies was composed, based on the results of the first experiment. Rather than include separate isotonic stimuli, we included stimuli that began with a series of repeated notes. The result was a $3 \times 2 \times 6$ factorial design: melodic direction (initial tones ascending, descending or repeated), articulation (legato vs. staccato), and tempo (with "L" element IOIs of 1200, 810, 650, 500, 428, and 375 ms), yielding 36 stimulus melodies in all. Rhythmic figure, interval size, and implied harmony were held constant (SSL, small intervals [stepwise motion], and tonic harmony, respectively). Staccato versions were made of each of sequence by taking the initial SSL rhythmic figures and dividing the "L" element into a short tone and a rest of equal length. This resulted in a figure whose absolute durations could be described as "SSSR" (with "R" indicating a rest whose duration was equal to the "S"); the IOIs for the tones in the melody were the same as in the legato versions. Appendix 2 gives legato versions of the three basic melodies, as well as a staccato version of the first melody.

The experiment used a within-subjects design; trials were conducted with individual participant sessions and sequences were presented in a quasi-random order. Stimuli were first randomized, and order of presentation manually adjusted to avoid consecutive presentations of either the same stimulus melody at different tempi or of runs longer than two stimuli at the same tempo. In addition, four foils (LSS melodies taken from the first experiment) were inserted in the block of stimuli to ensure that participants were sensitive to durational structure and provide variety over the course of the experimental session.

#### APPARATUS AND PROCEDURE

The apparatus and procedure were the same as in Experiment 1, save that the Macintosh G4 computer was connected to the Roland SPD6 and a Roland XV5050 sound module via a Digidesign 16 Channel MIDI interface. The "64voice Piano" tone setting was used. Participants listened to the stimuli over Beyerdynamic DT 770 headphones adjusted to a comfortable listening level.

### Results

As in Experiment 1, no significant effect of level of music training was observed, so data were analysed collapsed across participant groups. A repeated measures ANOVA with melodic direction (3 levels), articulation (2 levels) and tempo (6 levels) was conducted. As expected, there was a statistically significant main effect for tempo, $F(3.20, 57.63) = 6.47$, $p = .001$ (Greenhouse-Geisser correction applied); posthoc contrasts showed that slowest rate was significantly less upbeat (upbeatness index of .68) than all other levels (indices of .86, .92, .87, .91, and .88, respectively; $p = .007$, $\eta^2 = .34$). Differences in upbeatness between all other tempo levels were non-significant. A significant main effect for melodic direction was also found, $F(1.13, 20.38) = 7.44$, $p = .011$ (Greenhouse-Geisser correction applied). A posthoc contrast shows that this arises because both ascending (upbeatness index of .89) and descending (.90) melodies are significantly more upbeat than repeating tone melodies (index of .77; $p = .016/\eta^2 = .28$ and $p = .01/\eta^2 = .32$, respectively).

No main effect was found for articulation, and no interactions that involved it were found. In spite of this result, it was noted that repeated note sequences (Melody 3) were more strongly downbeat at the slowest tempo than were initially ascending or descending sequences, and possibly contributed disproportionately to the main effect for tempo; accordingly, data were reanalyzed excluding these sequences. The main effect for tempo remained, $F(2.66, 47.28) = 6.73$, $p = .001$, $\eta^2 = .27$ (Greenhouse-Geisser correction applied). Posthoc simple contrasts showed that the slowest tempo (upbeatness of .73) was almost significantly different from the next slowest level (upbeatness of .86; $p = .08$) but significantly different from all others (i.e., .95, .94, .94, .94; $p = .005$). The second-slowest level two is not significantly different from the faster tempo levels. With the repeated tone melodies excluded, there was no main effect for melodic direction, nor was there any main effect or interaction for articulation.

### Discussion

Experiment 2 gave a sharper focus to the effect of tempo on the upbeatness of a figure. While even at the slowest tempo these SSL melodies tended to be heard as anacrustic (index average of 0.68), there is a clear distinction between the slowest tempo versus all others. At that slowest tempo the IOI for the "S" elements is 600 ms; at the next slowest tempo the "S" IOI is 405 ms. Thus at the slowest tempo, the "S" IOIs themselves may project the primary beat or pulse (note that this 600 ms IOI is near the center of the range of "maximal pulse salience," Parncutt, 1994). Furthermore, a 600 ms IOI affords subdivision into shorter units, which again enhances its sense of beat (London, 2004, pp. 34-38). Thus, at the slowest tempo the SSL figure begins with "two beats," which readily invites a downbeat interpretation. At the next slowest tempo, the "S" elements are relatively fast, and indeed, are less likely candidates for further subdivision (see Repp, 2003). In addition, the reanalysis of the data shows that the presence of melodic motion enhances the effect of tempo on upbeatness, though from this design we were unable to discern any affect of direction, implied harmony, or other factors.

The lack of either a main effect or interaction that involved articulation was somewhat surprising, as articulation is one of the primary tools performers have to mark group boundaries and shape melodic gestures. While enhanced separation between groups may affect the expressive character of a melody, here it did not affect its accentual interpretation.

## Experiment 3

### *Method*

#### PARTICIPANTS

Participants were recruited from the same community as in Experiment 1, though with no overlap of participants between experiments. From an initial pool of 25 participants, 7 were excluded on the basis of not complying with task requirements; typically, either by tapping on every beat or being wholly unresponsive to variations in the stimuli by presenting the same pattern of responses to all stimuli, including foils that were included to test whether participants were in fact responding to stimulus structure (see discussions below). It is notable that participants were excluded from all groups—whether or not a participant was excluded did not depend on gender, age, or level of music training. Of the remaining 18 participants, 8 were male and 9 female; 17 were right-handed and 1 left-handed. Ages ranged from 19 to 59; the mean age was 36.8 years (*SD* = 16.1). In terms of musical background the same categories were used as in Experiments 1 and 2: less than 5 years of formal training (5 participants); 5-10 years (8 participants); and more than 10 years (5 participants). The average number of years of study was 7.4 years (*SD* = 5.2). Also as a result of the exclusions, order effects were not explored as participant exclusions meant that the number of participants per group per order was too low to form the basis for valid statistical inference.

### STIMULI AND DESIGN

To isolate the relative effects of melodic direction, implied harmony, and interval size, Experiment 3 involved the presentation of rhythmically isochronous melodies in a 2 × 2 × 2 × 2 factorial design: melodic direction (ascending vs. descending), implied harmony (tonic vs. dominant), interval size (skips vs. steps), and tempo (600ms vs. 460ms IOI). The full set of stimulus melodies used is given in Appendix 3. While exploring the effect of tempo was not a goal of this experiment, two tempos were used, primarily to provide some variety in the stimuli for our participants. Based on the first two experiments it was felt that at these tempos structural factors would have the most import, as they placed the note IOIs within a comfortable beat range (Parncutt, 1994).

### APPARATUS AND PROCEDURE

The apparatus and procedure were the same as in Experiment 1.

## Results

A mixed, repeated measures ANOVA was conducted on the results, with melodic direction (2 levels), implied harmony of initial note (2 levels), initial interval size (2 levels), and tempo (2 levels) as within-subjects variables, and training (3 levels) as a between-subjects variable. Responses to these isochronous melodies were overwhelmingly downbeat, as participants found most of the stimuli to be downbeat in most of the trials (grand average upbeatness value of .26). Nonetheless, there were a number of small but statistically significant main effects and interactions. There was a statistically significant main effect of melodic direction, with ascending melodies rated more upbeat than descending melodies (upbeatness index of .32 vs. .20; $F(1, 15) = 5.97$, $p = .03$, $\eta^2 = .29$; Greenhouse-Geisser correction applied). There was also a statistically significant main effect of implied harmony, as melodies that began with tones that implied a V chord (which then moved to tones that implied a tonic chord) were regarded as significantly more upbeat than those whose initial tones implied tonic harmony (upbeatness index of .31 vs. .21; $F(1, 15) = 9.51$, $p = .01$, $\eta^2 = .39$; Greenhouse-Geisser correction applied). As expected, there was no effect of tempo.

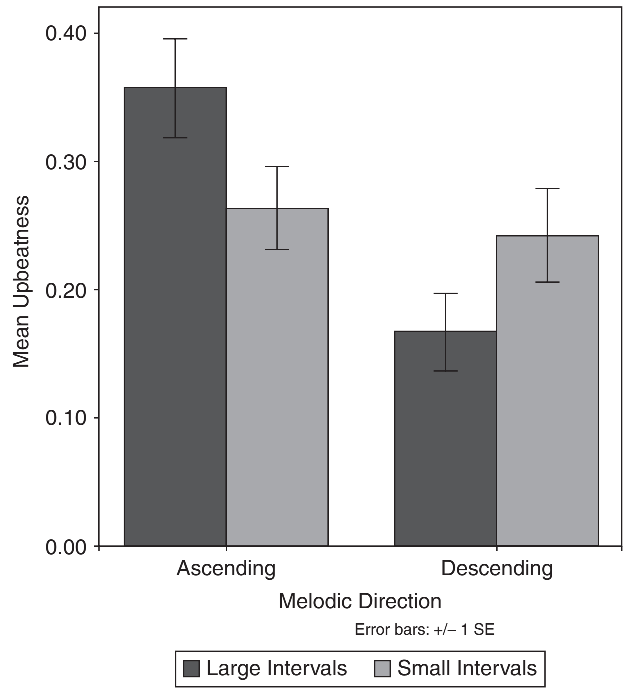


**FIGURE 5.** Interaction between direction and interval size, Experiment 3 (large Intervals > a major 3d; small intervals ≤ a major 3d).

There was a highly significant two-way interaction between direction and interval size, $F(1, 15) = 16.36$, $p = .001$, $\eta^2 = .52$ (Greenhouse-Geisser correction applied). Direction did not affect the upbeatness of small initial intervals (<M3), but larger initial intervals did show a dependence on melodic direction: ascending skips were far more upbeat than descending skips (see Figure 5).

Finally, in spite of the fact that there was no main effect of music training, there was a statistically significant three-way between-group and within-subjects interaction between implied harmony, interval size, and level of music training, $F(2, 15) = 15.02$, $p < .001$, $\eta^2 = .67$ (Greenhouse-Geisser correction applied). In essence, highly trained participants showed a greater sensitivity to stereotypical tonal cues (V-I motion occurring as a skip) and thus were more likely to respond to those sequences as upbeat (see Figure 6, panel c vs. panels a and b).

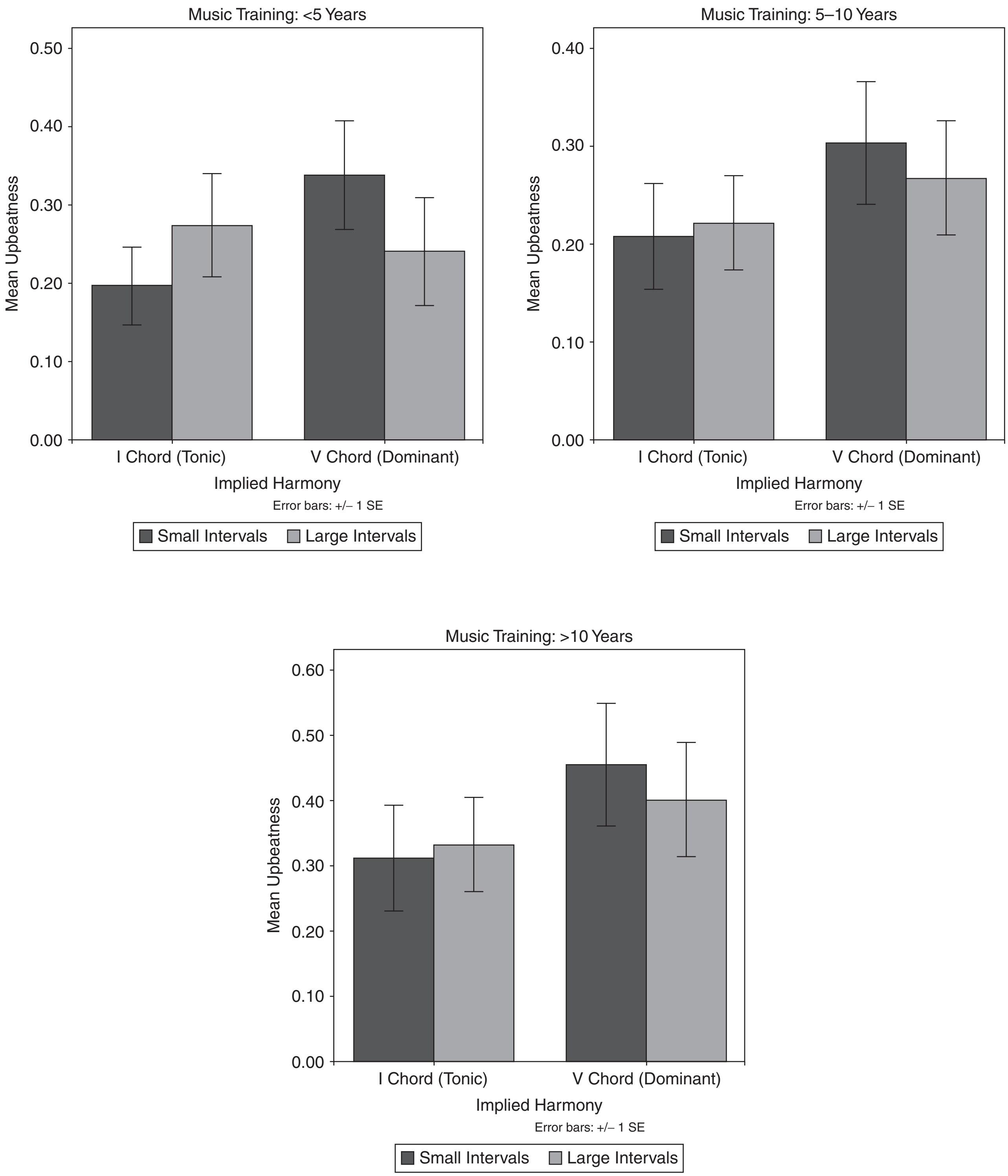


FIGURE 6. Interaction between implied harmony, interval size, and music training, Experiment 3.

## Discussion

While implied harmony does play a role in creating a sense of anacrusis, it is only a very weak marker of metric accent; most responses were non-anacrustic. Moreover, non-syntactic features of interval size and direction seem to play a role equal to syntactic cues (i.e., scale degree/implicit harmony) in creating a sense of upbeat. To put it another way, in these melodies unless a particular constellation of structural features was present (i.e., a combination of implied V-I harmony, large interval, and ascending motion), participants were unlikely to hear an anacrusis. Indeed, our finding that only in this experiment was music training relevant adds to the sense that the effect of tonal syntax on metrical accent is fairly limited. Only those participants with the highest level of music training (and hence a strong enculturation regarding the particular idioms of Western classical music) were likely to respond to those cues. Even this observation must be treated cautiously, for as noted above, some potential participants from this group were excluded due to their inability to perform the experimental task.

## General Discussion

The overall results of the three experiments are summarized in Table 1.

In these experiments durational structure was again shown to be the dominant factor in educing metrical responses. Other structural factors (initial melodic interval, initial melodic direction, and implied harmony) had a far more limited effect, as did performance factors (tempo and articulation). Indeed, articulation in the context of a rhythmic figure with well-defined durational structure seems to have had no effect whatsoever. This is in accord with previous studies, which examined the salience of relative IOI on determinations of accent and/or meter (e.g., Drake & Palmer, 1993; Lee, 1991; Povel & Essens, 1985; Steedman, 1977; and more recently, Toiviainen & Eerola, 2006)—in all these studies, relatively long IOIs are heard as relatively accented. This runs somewhat counter to the claims of Ellis and Jones (2009)—but see additional comments below. Likewise, Longuet-Higgins and Lee (1982), who specifically focused on the initial stages of beat (and hence accentual) induction, gave emphasis to the role of relatively long IOIs in generating a sense of beat and accent. Their main stratagem, however, is the avoidance of syncopations, which, like Steedman, involved aligning relatively long notes with the onset of beats and higher-level metrical units. The patterns used in our three experiments were designed to avoid problems of syncopation, as the duration of the two "S" elements was always equal to a single "L." But as Longuet-Higgins and Lee predicted, syncopation was wholly avoided, as none of our participants ever construed a pattern in terms of a syncope (e.g., hearing the second "S" as being on the downbeat: S|SLSSLS|SLS . . .). Longuet-Higgins and Lee (1984) also claimed that many melodies are metrically ambiguous. While we found this to be true, we found that the initial rhythmic configuration strongly biases the metrical percept. To be sure, we were careful not to present stimuli that were ambiguous as to the duration of the basic beat (e.g., a 3/4 vs. 6/8 interpretation of a series of eighth notes), or the structure of the measure (2/4 vs. 4/4). Nonetheless, in our experiments we found that melodies that begin with an LSS figure were overwhelmingly downbeat, and the manipulation of structural and/or performance factors seems to have little effect on their tendency to be heard as such. SSL rhythms, on the other hand, while strongly tending toward anacrustic interpretation, were more flexible in their metrical interpretation. As such, the sense of anacrusis in a melody that starts with an SSL figure may be manipulated by both structural and performance factors. Thus, rhythmic structure, while central, is not the whole story.

Moreover, when isotonic melodies were compared with non-isotonic (i.e., ascending or descending) patterns, the result is that the addition of melodic structure of any sort rendered both LSS and SSL stimuli more upbeat by comparison. Thus, there seems to be a broad interaction between rhythm and all other structural factors. To put it another way, if a melody lacks a change of pitch, not much "moves" in a phenomenological sense. The very terms "upbeat" and "downbeat" are imbued with a sense of motion, and whether this motion occurs in a "virtual pitch space" (Langer, 1953), or is "acousmatic" (Scruton, 1997), or is a kind of perceptual illusion (Gjerdingen, 1994) does not matter. The phenomenological upshot is that patterns that involve tonal differentiation may be qualitatively different from those that do not. Thus we have been cautious when making comparison(s) between these two sets of stimuli and our participants' responses to them.

Experiment 2 showed that tempo can effect upbeatness of SSL melodies, and thus upbeatness is not just an aspect of durational pattern (i.e., the ratios of successive IOIs), but is also due to the absolute value(s) of the durations themselves. To be sure, this tempo effect occurred only at the slowest level, though there seem to be vestiges of it at the next-slowest level. This aligns with the categorical boundary (around 400 ms) between "temps longs" and "temps courts" suggested by Fraisse (1963), and also with London's (2004) hypothesis that to hear an interval as a beat requires the possibility of subdivision (if such subdivision is not explicitly present in the melody itself or elsewhere in the musical texture). Thus the tempo-related shift for upbeat vs. downbeat SSL melodies occurred at the point where the "S" elements move from IOIs that were short enough to be heard as subdivisions to IOIs that became "subdivide-able" beats in their own right. The tempo effect observed in Experiment 2 also gives credence to London's notion of tempo-metrical types, based on the relationship between the periodicities in certain ranges (roughly100-400 ms, 400-1200 ms, 1200-2000 ms) that are evident in a given metrical context (see London, 2004, pp. 76-79, as well as Repp, London, & Keller, 2005, 2008).

Our results reinforce that Huron and Royal's (1996) previously mentioned scepticism regarding the relative influence of melodic accent. In all three experiments an effect of melodic direction was observed, though it was a slight main effect. Only in Experiment 3, with rhythmic structure removed, did harmonic motion and interval size have a statistically significant effect. Yet even then, the effect of those factors on upbeatness was similarly modest; for the most part, isochronous melodies are heard as downbeat, rather than anacrustic. An upbeat perception in a rhythmically isochronous context is more likely if the listener is a highly trained musician and thus familiar with the idiomatic "pickup note" gestures that are commonplaces in Baroque, Classical, and Romantic music. Harmony alone (which may be abstracted from any particular configuration of pitch chroma or height), however, is unlikely to create a sense of anacrusis for most listeners. This finding runs counter to the claims of much music theory, which posits that harmonic structure is the primary determinant of metric accent. As the interactions noted in Table 1 show, the more concrete structural aspects of melodic direction and interval size are of equal or greater efficacy in creating a sense of upbeat. Our findings also run counter, somewhat, to the claims of Ellis and Jones (2009) regarding "joint accent structure" (see also Jones & Boltz, 1989; Jones & Pfordresher, 1997). Ellis and Jones argue that a combination of melodic (pitch-based) and temporal (rhythmic, based on tone duration/articulation) accents is involved in metric perception, and our results support this claim. They also argue against the "temporal accent bias hypothesis" (p. 265), which claims that temporal factors play the primary role in meter perception. However, our results support the temporal accent bias hypothesis, as durational structure (SSL vs. LSS) was the primary determinant of anacrustic versus non-anacrustic perception.[1] More broadly, we would argue that joint accent structure and temporal bias are not mutually exclusive; like Ellis and Jones, we note that the relative salience of each needs to be carefully studied using controlled stimuli, such as those employed in our experiments.

The metric ambiguity of our stimuli also gave rise to problems in Experiment 3, where, as noted above, the data from a good number of participants had to be rejected. Absent of any rhythmic structure that made the presence of an upbeat or downbeat clear (or at least plausible), we observed that a number of participants would consistently start tapping on the third or fourth tone, regardless of stimulus structure. Thus, rather than testing for sensitivity to upbeat/downbeat cues, these participants used one or two IOIs to gauge the tempo, and then started tapping. These participants (whom we nicknamed either the "doggedly downbeat" or "assiduously upbeat" tappers) seemed to treat the experiment as a kind of tempo-tracking and synchronization task. For these participants the structural configuration of the melodies (and the rhythmic foils!) had no metrical effect whatsoever. While this data had to be rejected, it was another indication of the relatively weak effect that structural factors have on upbeatness in isochronous melodies at moderate tempi.

Finally, there are two larger implications from our experiments. First, Experiment 1 (and to a lesser extent, Experiment 2) demonstrated that rhythms based on repeated tones (i.e., isotonic sequences) do not behave in the same way as sequences with more complex melodic structure(s). Likewise, Experiment 3 demonstrated that isochronous sequences do not behave in the same way as rhythmically differentiated sequences. In both cases, there seems to be a complexity threshold or divide which phenomenally separates these different classes of stimuli. This also suggests, to use the terms of Ellis and Jones (2009), that the interactions between rhythmic and melodic factors are "interactive" rather than "additive"—the contributions of interval size, direction, harmony, rhythmic figure, and articulation

[1]Ellis and Jones (2009) examined the effect of duration using stimuli with isochronous onset intervals (e.g., rhythms involving a long-short-rest, where the duration of the short-rest was equal to the long). As such, they did not investigate joint accent structure that involved durationally marked temporal accents, as we have done here.

must be analyzed in context. Caution must be taken, therefore, in generalizing from studies of isochronous, melodically-undifferentiated stimuli to more ecologically-valid musical contexts. The second implication is that "tempo effects are complex." As Experiment 1 demonstrated, one simply cannot say that faster tempos make rhythms more anacrustic; changes in tempo interact with rhythmic structure, and though our experiments did not address these issues, it is likely that tempo affects the perception of melodic interval and direction as well ("fast" versus "slow" leaps having different gestural connotations). The effect of global tempo on the melodic, rhythmic, metric, and expressive aspects of music would thus seem to be an area ripe for further study.

## Author Note

Special thanks to the choristers of Wolfson College for their participation in Experiment 2.

This research was supported by a Fulbright UK Distinguished Scholars Lecturing/Research Grant (Justin London) and by the Helsingin Sanomat Centennial foundation (Tommi Himberg). The authors wish to thank Devin McAuley, Stephen Handel, and one anonymous reviewer for their very helpful comments on the manuscript.

*Correspondence concerning this article should be addressed to* Justin London, Department of Music, Carleton College, One North College St., Northfield, MN 55057. E-MAIL: jlondon@carleton.edu

## References

BENJAMIN, W. E. (1984). A theory of musical meter. *Music Perception, 1,* 355-413.

CONE, E. T. (1968). *Musical form and musical performance*. New York: Norton.

COOPER, G., & MEYER, L. B. (1960). *The rhythmic structure of music*. Chicago, IL: University of Chicago Press.

DESAIN, P., & HONING, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research, 28,* 29-42.

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

DRAKE, C., & PALMER, C. (1993). Accent structures in musical performance. *Music Perception, 10,* 343-378.

ELLIS, R. J., & JONES, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 264-280.

FRAISSE, P. (1963). *Psychology of time*. New York: Harper.

GJERDINGEN, R. O. (1994). Apparent motion in music? *Music Perception, 11,* 335-370.

HANDEL, S. (1998). The interplay between metric and figural organization. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1546-1561.

HASTY, C. (1997). *Meter as rhythm*. Oxford, UK: Oxford University Press.

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

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

JONES, M. R., & PFORDRESHER, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology, 51,* 271-291.

KRAMER, J. D. (1988). *The time of music.* New York: Schirmer Books.

KRUMHANSL, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126,* 159-79.

LANGER, S. (1953). *Feeling and form.* New York: Scribners.

LEE, C. S. (1991). The perception of metrical structure: Experimental evidence and a model. In P. Howell, R. West, & I. Cross (Eds.), *Representing musical structure* (pp. 59-127). New York: Academic Press.

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

LONDON, J. M. (2001). Rhythm. In J. Tyrell & S. Sadie (Eds.), *The new Grove dictionary of music and musicians* (Vol. 21, pp. 277-309). Oxford, UK: Oxford University Press.

LONDON, J. M. (2004). *Hearing in time.* New York: Oxford University Press.

LONGUET-HIGGINS, H. C., & LEE, C. S. (1982). The perception of musical rhythms. *Perception, 11,* 115-128.

LONGUET-HIGGINS, H. C., & LEE, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception*, *1,* 424-441.

NARMOUR, E. (1990). *The analysis and cognition of basic melodic structures.* Chicago, IL: The University of Chicago Press.

PARNCUTT, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception, 11,* 409-464.

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

REPP, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, 35, 355-370.

Repp, B. H., Iverson, J. R., & Patel, A. D. (2008). Tracking an imposed beat within a metrical grid. *Music Perception, 26,* 1-18.

Repp, B. H., London, J., & Keller, P. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Perception*, *23*, 61-78.

Repp, B. H., London, J., & Keller, P. (2008). Phase correction in sensorimotor synchronization with non-isochronous rhythms. *Music Perception, 26,* 171-175.

Rothstein, W. N. (1989). *Phrase rhythm in tonal music*. New York: Schirmer Books.

Scruton, R. (1997). *The aesthetics of music*. Oxford, UK: Oxford University Press.

Sloboda, J. A. (1983). The communication of musical metre in piano performance. *Quarterly Journal of Experimental Psychology, 35A,* 377-396.

Steedman, M. J. (1977). The perception of musical rhythm and metre. *Perception*, *6,* 555-569.

Toiviainen, P., & Eerola, T. (2003). Where is the beat?: Comparison of Finnish and South African listeners. In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), *Proceedings of the 5th Triennial ESCOM Conference* (pp. 501-504). Hanover, Germany: Hanover University of Music and Drama. link:http://users.jyu.fi/~ptee/publications/1_2003.pdf

Toiviainen, P., & Eerola, T. (2006). Autocorrelation in meter induction: The role of accent structure. *Journal of the Acoustical Society of America, 119,* 1164-1170.

Yeston, M. (1976). *The stratification of musical rhythm*. New Haven: Yale University Press.

# Appendices

## Experiment 1 Stimuli

## Experiment 2 Stimuli

## Experiment 3 Stimuli