Published in final edited form as:

J Exp Psychol Hum Percept Perform. 2013 August ; 39(4): . doi:10.1037/a0031087.

Temporal Dynamics and the Identification of Musical Key

Morwaread Mary Farbood,

Department of Music and Performing Arts Professions, Steinhardt School, New York University

Gary Marcus, and

Department of Psychology, New York University

David Poeppel

Department of Psychology, Center for Neural Science, New York University

Abstract

A central process in music cognition involves the identification of key, however little is known about how listeners accomplish this task in real time. This study derives from work that suggests overlap between the neural and cognitive resources underlying the analyses of both music and speech, and is the first to explore the timescales at which the brain infers musical key. We investigated the temporal psychophysics of key-finding over a wide range of tempi using melodic sequences with strong structural cues, where statistical information about overall key profile was ambiguous. Listeners were able to provide robust judgments within specific limits, at rates as high as 400 beats per minute (~7 Hz) and as low as 30 bpm (0.5 Hz), but not outside those bounds. These boundaries on reliable performance show that the process of key-finding is restricted to timescales that are closely aligned with beat induction and speech processing.

Keywords

music perception; tonal induction; temporal processing; speech; rate

Speech and music, two of the most sophisticated forms of human expression, differ in fundamental ways. Although hierarchical elements of music such as harmony have been argued to resemble syntactic structures in language, these structures do not have semantic content in the sense conveyed by language (Slevc & Patel, 2011). Discrete pitch, one of the basic units of musical structure, is not utilized in speech. Although continuous pitch change is an aspect of intonation, the building blocks of speech are encoded primarily through timbral changes (Patel, 2008; Zatorre, Belin, & Penhune, 2002). Furthermore, music has a vertical (harmonic) dimension and a rhythmic-metrical aspect that are both absent in speech. Nonetheless, music and speech are both highly structured, complex auditory signals, and an important question is whether there is significant overlap in the neurocomputational resources that form the basis for processing both types of signals. The motivation for this study derives in part from recent work that suggests overlap between the neural and cognitive resources underlying the structural processing of both music and language (Carrus, Koelsch, & Bhattacharya, 2011; Ettliger, Margulis, & Wong, 2011; Fedorenko, Patel, Casasanto, Winawer & Gibson, 2009; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Kraus & Chandrasekaran, 2010; Patel, 2008). While the majority of previous work has explored higher-level *cognitive* aspects of music and language—in particular shared resources for

syntactic processing—the present study is focused on the *timescales* at which the brain infers musical key and how they compare to timescales implicated in speech.

Because the modulation spectra of speech and music have similar peaks (ranging from 2–8 Hz), it seems plausible that both are parsed and decoded at comparable rates. Melodies, like spoken sentences, consist of patterns of sound structured in time. To understand a sentence, a listener must recover the features, (di)phones, syllables, words, and phrases that form a sentence's constituent parts. Perhaps the closest musical analog to speech comprehension is key-finding, which involves the perception of hierarchical relationships between notes and intervals and how they are interpreted in a larger context. Identification of a tonal center is a process that is at the core of how all listeners experience music, yet little is known about how such inferences are derived in real time.

The most prominently debated theory of musical key recognition is premised on the idea that listeners extract zeroth-order statistical distributions of the pitch classes in a piece and then identify key based on the degree to which those distributions correlate with prototypical distributions (“key profiles”) (Krumhansl & Kessler, 1982; Krumhansl, 1990; Longuet-Higgins & Steedman, 1971; Temperley, 2007; Vos & Van Geenen, 1996; Yoshino & Abe, 2004). However, other work has indicated that purely statistical approaches do not offer a complete account of how listeners identify key, suggesting that key recognition involves structural factors (Brown, 1988; Brown, Butler, & Jones, 1994; Butler, 1989; Matsunaga & Abe, 2005; Temperley & Marvin, 2008; Vos, 1999). In essence, zeroth-order statistical distributions might be an epiphenomenon that falls out of the melodic structural schemas that are essential to the recognition of a tonal center. In light of these concerns, our exploration of the temporal psychophysics of key-finding focused on musical stimuli that contained identical pitch material prior to transposition.

A useful dichotomy for categorizing key-finding approaches is the distinction between bottom-up and top-down processing (Parcutt & Bregman, 2000). Bottom-up processing depends on information drawn directly from the stimuli, reflecting the influence of immediately preceding pitches in short-term or sensory memory. Top-down processing is based on schemata that are activated from long-term memory and applied to a musical passage by the listener. Bottom-up approaches to modeling key-finding have been employed less frequently and are often combined with top-down frameworks. One such example is Huron and Parcutt's (1993) method, which extended Krumhansl's (1990) key profile approach by taking into account psychoacoustic factors and sensory memory decay. Although these modifications improved the model predictions, it still failed to account for Brown's (1988) experimental findings regarding the importance of intervallic structure for melodic key-finding. Leman's (2000) model, based on echoic images of periodicity pitch, is an example of a purely bottom-up approach. Leman challenges the claim that tonal induction in probe-tone experiments is based on top-down processing. However, he cautions that although his model appears to model degree of fitness for a probe tone in a tonal context successfully, a schema-based model is still required for actual *recognition* of a tonal center.

Harmonic priming studies have illuminated the contributions of both cognitive (top-down) and sensory (bottom-up) processing. In general, these studies have found that a chord is processed faster in a harmonically related context than an unrelated context (Bharucha, 1987; Bharucha & Stoeckig, 1986; Bharucha & Stoeckig, 1987; Bigand & Pineau, 1997; Tillmann & Bigand, 2001; Tillmann, Bigand, & Pineau, 1998), and that both sensory and cognitive components are involved in musical priming (Bigand, Poulin, Tillmann, Madurell, & D'Adamo, 2003; Tekman & Bharucha, 1998). Bigand et al. (2003) observed that cognitive priming systematically overruled sensory priming except at the fastest tempo they explored (75 ms per chord). This indicates that while key-finding can be accomplished

rapidly, there still exists a rate limit. Discovering the boundaries of this limit and comparing them to known timescales implicated in speech processing are the primary goals of this study.

Experiment 1

Method

Experiment 1 was the initial study in which we obtained key labels for our statistically “neutral” stimuli. A subset of these stimuli were then used in Experiment 2, the main experiment, in which we assessed the time course over which listeners make robust key judgments. For Experiment 1, we constructed 31 eight-note melodic sequences that fell into three structural categories: two types had *strong structural cues* intended to invoke one of two possible keys, and the third contained *little or no structural cues*.

The starting point for constructing our materials was the fact that keys that differ by only one sharp or flat overlap almost completely in their sets of underlying notes. The union of the two such keys, C major and G major, consists of C, D, E, F, F#, G, A, B, a set of pitches that is inherently ambiguous between the two keys. Our experiments explored permutations of these statistically ambiguous collections of notes. For expository purposes, we will refer to the two keys as “lower” (C major) and “upper” (G major). Several music-theoretic guidelines were used to compose melodies with strong structural cues:

- Tendency tones—pitches in a particular key that are commonly followed by another pitch within that key—were resolved.¹
- The contour of the pitches clearly outlined common chords in Western harmony.²
- Chords implied by the ordering of the pitches frequently followed syntactically predictable progressions.³

We controlled for the effect of recency on short-term memory by ensuring that all sequences ended on the same note, the tonic of the upper key (e.g., G in the case of C/G major). In addition, we constrained the penultimate note to always be either a second or a third above the final note; these two ending types were distributed evenly among the sequences. In this way, the final note functioned in every trial as a musically critical note, regardless of which key a listener inferred. All 31 sequences consisted of monophonic, isochronous tones rendered in a MIDI grand piano timbre. The inter-onset interval between note events was 600 ms and the sequences were randomly transposed to all 12 chromatic pitch class levels. There were 10 sequences in each of the two key categories, and 11 in the ambiguous category.⁴

Participants and Task—Six experts with professional-level training in music theory participated. The subjects accessed the study through a website that presented the 31 melodic sequences in pseudorandom order. In addition to the audio playback, each sequence was accompanied by a visual representation in staff notation. Participants were asked to specify the key for each melody; if they felt that the sequence was not in any particular key,

¹For the ambiguous sequences, tendency tones were subverted. For example, possible leading tones in both the upper and lower keys (tones that are expected to resolve half a step up to a tonic) were placed after the resolving tone, in a different register than the resolving tone, or temporally distant from the resolving tone.

²Typical chords outlined included I, V⁷, IV, ii.

³In particular, a subdominant-dominant-tonic progression was outlined for upper key sequences and a tonic-dominant-tonic progression for lower key sequences.

⁴There were originally 10 ambiguous sequences to match the 10 in the other two categories, but one more was added to test the assumption that a clearly outlined, syntactically unexpected progression would result in ambiguous key perception.

they were instructed to label it “ambiguous.” Additionally, they were asked to rate the confidence of their response on a scale from 1 to 4 (1 = very unsure, 4 = very confident).

Results

The complete set of stimuli and data are provided in the Appendix (Table A1). Ratings were quantified by assigning negative values to lower key responses and positive values to upper key responses with magnitudes corresponding to the confidence values. Ambiguous responses were assigned a value of 0. Consistent with predictions derived from music-theoretic principles, structural factors determined listeners’ judgment of key despite the ambiguous statistical profiles. Melodic sequences that were predicted to be perceived as belonging to the lower key received a within-subject average rating of -2.42 ($SD = 0.95$), while sequences predicted to belong to the upper key received a mean rating of 1.85 ($SD = 2.04$), with passages predicted as ambiguous receiving intermediate responses (mean 0.09 , $SD = 1.13$), $F(2, 10) = 17.48$, $p = 0.0005$. Post-hoc Tukey-Kramer tests revealed that the upper and lower key categories differed significantly from each other, and that the lower key category differed significantly from the ambiguous category as well (the type of ending, descending major second versus major third from the penultimate to the final note, was not correlated with overall rating, $t(184) = -0.67$, $p = 0.50$). Figure 1 shows the five sequences most clearly eliciting the lower and upper keys. These 10 sequences served as the materials for the main experiment.

Experiment 2

Method

Participants—The participants were 22 university students (mean age 23.8 years; 14 male) who were skilled at instrumental performance and had an average of 15.5 years of musical training ($SD = 6.4$) and had taken at least one music theory course. Two additional subjects, self-rated a 2 or lower on an overall musical proficiency scale of 1 (lowest) to 5 (highest), were excluded because they could not execute the task, presumably due to lack of sufficient musical training.

Materials—Each of the 10 sequences depicted in Figure 1 were rendered in MIDI grand piano timbre at 7, 15, 30, 45, 60, 75, 95, 120, 200, 400, 600, 800, 1000, 1200, 1600, 2200, and 3400 bpm, although the first five subjects were not exposed to the sequences at 3400 bpm.

Task—Participants were presented with one sequence per trial on Sony MDR-CD180 headphones and asked to indicate whether each sequence sounded resolved (ending on an implied tonic) or unresolved (ending on an implied dominant) by entering responses into a Matlab GUI that used Psychtoolbox for audio playback. Subjects were instructed to ignore aspects such as perceived rhythmic or metrical stability when making their decision.

Each participant listened to 170 sequences (160 for the initial five subjects) in a pseudorandomized order that took into account tempo, key, and original sequence, such that no stimulus was preceded by another stimulus generated from the same original sequence or having the same tempo, and no stimulus was in the same key as the *two* preceding stimuli. All stimuli were transposed such that they were at least three sharps/flats away from the key of the immediately preceding stimulus.

Results

Figure 2 (bottom panel) shows the mean percent correct responses as well as d' values for each tempo across all sequences and all subjects. Visual inspection of the psychophysical

data reveals a performance plateau, with a preferred range of tempi in which participants provide the most robust judgments, from approximately 30–400 bpm. Judgment consistency sharply decreases for tempi below 30 bpm and above 400 bpm, with a fairly steep decline occurring above 400. A one-way, repeated-measures ANOVA, excluding the initial five subjects who were not exposed to the 3400 bpm case, revealed a significant effect of tempo, $F(5.87, 93.92) = 20.61, p < .001$ (Greenhouse-Geisser corrected). Post-hoc multiple comparisons performed using Tukey's HSD test (Table 1), supported by quadratic trend contrasts, $F(1, 331) = 162.53, p < .001$, indicate that accuracy was significantly greater for tempi within the 30–400 bpm temporal zone than for tempi outside that zone (7–15, 600–3400 bpm).

Discussion

The findings provide a new perspective on how musical knowledge is deployed online in the determination of a tonal center or key. In Experiment 1, expert listeners categorized materials that were constructed to be statistically ambiguous, thus requiring classification based on structural cues. We utilized these stimuli in Experiment 2, where we observed an inverted U-shaped curve with a temporal “sweet spot” for analyzing an input sequence and being able to determine its tonal center: between 30–400 bpm (0.5–6.7 Hz modulation frequency; 2 s to 150 ms IOI). Listeners were highly consistent in their structurally cued classification and remarkably quick in inferring a tonal center for a sequence, capable of reliably identifying the key after just seven notes presented within 1.05 seconds. Our data thus (i) support the existence and utility of abstract, structural information in the perceptual analysis and processing of music and (ii) show the extent to which it is integrated into processing systems with particular temporal resolution and integration thresholds.

The results point to clear processing constraints, both at high and low stimulus rates. At the high rate (400 bpm), listeners require ~150 ms per note to generate the response profile observed. Although elementary auditory phenomena such as pitch detection, order threshold, and frequency modulation direction detection are associated with much shorter time constants (~20–40 ms; see Divenyi, 2004; Hirsh, 1959; Warren, 2008; White & Plack, 1998), the longer time course we identify for the aggregation of structural information in key-finding implicates the need for extra processing time for extracting melodic structure.

At rates below about 30 bpm, the sequences apparently fail to integrate into perceptual objects that permit the relevant operations. Presumably, the interaction of the temporal integration and working memory mechanisms that jointly underlie the construction of objects of a suitable granularity are increasingly challenged at slower rates. Our data provide a numerical confirmation of studies by Warren, Gardner, Brubaker, & Bashford (1991) who used very different materials to test the recognition of known melodies and found ~150 ms lower to ~2000 ms upper bounds for their task.

From a note-event perspective, the temporal range over which key-finding is optimal is similar though not identical to critical time constants implicated in processing continuous speech. The modulation frequencies over which speech intelligibility is best ranges from ~2–10 Hz (delta and theta bands) (Ghitza, 2011; Giraud et al., 2000; Luo & Poeppel, 2007). These numbers align with the peak of the modulation spectrum of speech, which across languages tends to lie between 4–6 Hz (Greenberg, 2006). In the melodic sequence case examined here, the ideal range is a bit lower, with optimal performance centered in the low delta to low theta range (0.5–6 Hz). Notably, this also aligns very closely with the typical range (30–300 bpm/0.5–5 Hz/50–2000 ms IOI) in which listeners can detect rhythmic pulse (with a preferred pulse of around 100 bpm/1.7 Hz/588 ms IOI) (London, 2004). Beat induction and key-finding presumably represent very different processes, but both are

foundational to music. The very close alignment of these two ranges seems to imply that both processes are limited by the same mechanisms.

Figure 2 (top panel) presents a comparison of various processing thresholds for both music and speech and depicts how the data from the main experiment align with them. The findings underscore both principled similarities between the two domains in the overall temporal processing range—consistent with hypotheses about shared resources—as well as specific differences (peaks at ~2 Hz versus ~5 Hz), arguably attributable to the different representations or data structures that form the basis of music versus speech.

A significant difference between the two domains is the presence of a vertical dimension in the form of chords and harmony in music. The fact that this dimension is not utilized in our monophonic stimuli arguably increased the difficulty of the key-finding task. It can be further argued that the stimuli constructed for this study are not representative of normal music and that key identification would actually happen much faster if the pitch profiles were not ambiguous and chords were present. However, findings from priming studies do not support this. In particular, Bigand et al.'s (2003) study comparing sensory versus cognitive components in harmonic priming offers another perspective on tonal induction at fast tempi. The stimuli for this study consisted of eight-chord sequences in which the first seven chords served as a context for a final target chord (paralleling the eight-note structure of the melodies here). They found that at 300 and 150 ms per chord, the “cognitive component” clearly facilitated processing of the target, indicating that key-finding had successfully occurred despite the very fast tempo. However, when the tempo was further increased to 75 ms per chord (800 bpm/13.3 Hz), the cognitive component was marginal for musicians and seemingly overruled by the sensory component for nonmusicians. This marked difference between the 150 and 75 ms cases aligns closely with the current data and indicates that regardless of the information content, there is a minimum amount of processing time that is necessary for key induction.

Although we used expert listeners in our pilot study and musically experienced listeners in our main study, they provide a window into a *universal* process; just as language is universal to all speakers, key-finding is universal to all listeners, whether musically trained or not (see Bigand & Poulin-Charronnat, 2006 for review). Our results provide principled bounds on the rates at which structure can be integrated into the process of key-finding and speak to both the subtle differences and similarities in how music and speech are processed. While each system presumably relies on its own proprietary database of constituent elements (e.g. phonemes, syllables, and words for languages, motivic-intervallic elements for music), common physiological properties place broad constraints on the mechanisms by which humans listeners can decode streams of auditory information, whether linguistic, musical, or otherwise.

Acknowledgments

We thank Ran Liu, Josh McDermott, and David Temperley for critical comments on the manuscript. This work is supported by NIH 2R01 05660 awarded to DP.

References

- Bharucha JJ. Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*. 1987; 5:1–30.
- Bharucha JJ, Stoeckig K. Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*. 1986; 12:403–410. [PubMed: 2946797]

- Bharucha JJ, Stoeckig K. Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics*. 1987; 41:519–524. [PubMed: 3615148]
- Bigand E, Pineau M. Global context effects on musical expectancy. *Perception & Psychophysics*. 1997; 59:1098–1107. [PubMed: 9360482]
- Bigand E, Poulin-Charronnat B. Are we “experienced listeners?” A review of the musical capacities that do not depend on formal musical training. *Cognition*. 2006; 100:100–130. [PubMed: 16412412]
- Bigand E, Poulin B, Tillmann B, Madurell F, D'Adamo DA. Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*. 2003; 29:159–171. [PubMed: 12669755]
- Brown H. The Interplay of Set Content and Temporal Context in a Functional Theory of Tonality Perception. *Music Perception*. 1988; 5:219–250.
- Brown H, Butler D, Jones MR. Musical and temporal influences on key discovery. *Music Perception*. 1994; 11:371–407.
- Butler D. Describing the Perception of Tonality in Music: A Critique of the Tonal Hierarchy Theory and a Proposal for a Theory of Intervallic Rivalry. *Music Perception*. 1989; 6:219–242.
- Carrus E, Koelsch S, Bhattacharya J. Shadows of music-language interaction on low frequency brain oscillatory patterns. *Brain and Language*. 2011; 119:50–57. [PubMed: 21683995]
- Divenyi PL. The times of Ira Hirsh: Multiple ranges of auditory temporal perception. *Seminars in Hearing*. 2004; 25:229–239. [PubMed: 16479266]
- Ettlinger M, Margulis EH, Wong PCM. Implicit Memory in Music and Language. *Frontiers in Psychology*. 2011; 2:1–10. [PubMed: 21713130]
- Fedorenko E, Patel A, Casasanto D, Winawer J, Gibson E. Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*. 2009; 37:1–9. [PubMed: 19103970]
- Ghitza O. Linking Speech Perception and Neurophysiology: Speech Decoding Guided by Cascaded Oscillators Locked to the Input Rhythm. *Frontiers in Psychology*. 2011; 2:1–13. [PubMed: 21713130]
- Giraud A-L, Lorenzi C, Ashburner J, Wable J, Johsrude I, Frackowiak R, Kleinschmidt A. Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology*. 2000; 84:1588–1598. [PubMed: 10980029]
- Greenberg, S. A Multi-tier framework for understanding spoken language. In: Greenberg, S.; Ainsworth, W., editors. *Listening to Speech: An Auditory Perspective*. 2006. p. 1-32.
- Hirsh IJ. Auditory perception of temporal order. *Journal of the Acoustical Society of America*. 1959; 31:759–767.
- Huron D, Parncutt R. An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology*. 1993; 12:154–171.
- Koelsch S, Gunter TC, Wittfoth M, Sammler D. Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*. 2005; 17:1565–1577. [PubMed: 16269097]
- Kraus N, Chandrasekaran B. Music training for the development of auditory skills. *Nature Reviews Neuroscience*. 2010; 11:599–605.
- Krumhansl, CL. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press; 1990.
- Krumhansl CL, Kessler EJ. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*. 1982; 89:334–368. [PubMed: 7134332]
- Leman M. An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*. 2000; 17:481–509.
- London, J. *Hearing in Time: Psychological Aspects of Musical Meter*. New York: Oxford University Press; 2004.
- Longuet-Higgins HC, Steedman MJ. On interpreting Bach. *Machine Intelligence*. 1971; 6:221–241.
- Luo H, Poeppel D. Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*. 2007; 54:1001–1010. [PubMed: 17582338]
- Matsunaga R, Abe J. Cues for Key Perception of a Melody. *Music Perception*. 2005; 23:153–164.
- Parncutt R, Bregman AS. Tone profiles following short chord progressions: Top-down or bottom-up? *Music Perception*. 2000; 18:25–57.

- Patel, A. Music, Language, and the Brain. New York: Oxford University Press; 2008.
- Slevc LR, Patel AD. Meaning in music and language: Three key differences. Comment on “Towards a neural basis of processing musical semantics” by Stefan Koelsch. *Physics of Life Reviews*. 2011; 8(2):110–111. [PubMed: 21570367]
- Tekman HG, Bharucha JJ. Implicit knowledge versus psychoacoustic similarity in priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*. 1998; 24:252–260.
- Temperley, D. Music and Probability. Cambridge, MA: MIT Press; 2007.
- Temperley D, Marvin EW. Pitch-class distribution and the identification of key. *Music Perception*. 2008; 25:193–212.
- Tillmann B, Bigand E. Global context effect in normal and scrambled musical sequences. *Journal of Experimental Psychology: Human Perception and Performance*. 2001; 27:1185–1196. [PubMed: 11642702]
- Tillmann B, Bigand E, Pineau M. Effects of global and local contexts on harmonic expectancy. *Music Perception*. 1998; 16:99–117.
- Vos PG. Key implications of ascending fourth and descending fifth openings. *Psychology of Music*. 1999; 27:4–17.
- Vos PG, Van Geenen EW. A parallel-processing key-finding model. *Music Perception*. 1996; 14:185–223.
- Warren, RM. Auditory Perception: An Analysis and Synthesis. 3rd ed.. Cambridge, UK: Cambridge University Press; 2008.
- Warren RM, Gardner DA, Brubaker BS, Bashford JA. Melodic and nonmelodic sequences of tones: Effects of duration on perception. *Music Perception*. 1991; 8:277–289.
- White LJ, Plack CJ. Temporal processing of the pitch of complex tones. *Journal of the Acoustical Society of America*. 1998; 108:2051–2063. [PubMed: 9566327]
- Yoshino I, Abe J-I. Cognitive modeling of key interpretation in melody perception. *Japanese Psychological Research*. 2004; 46(4):283–297.
- Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*. 2002; 6:37–46. [PubMed: 11849614]

Appendix

Table A1

Complete results for Experiment 1.

Predicted key	Stim. Num	Melodic sequence	Ending type	Mean Score	Std. dev.
Lower key	16		M2	−3.00	0.00
Lower key	20		M3	−3.00	0.71
Lower key	3		M3	−2.80	0.45
Lower key	7		M2	−2.60	1.52

Predicted key	Stim. Num	Melodic sequence	Ending type	Mean Score	Std. dev.
Lower key	27		M3	-2.60	0.89
Lower key	11		M2	-2.20	1.30
Lower key	30		M3	-2.00	2.00
Lower key	12		M2	-1.60	1.34
Ambiguous	31		M3	-1.60	2.51
Lower key	22		M2	-1.60	2.88
Lower key	23		M3	-1.20	1.30
Lower key	4		M2	-1.00	3.74
Ambiguous	26		M2	-0.80	1.92
Ambiguous	18		M3	-0.60	1.95
Ambiguous	13		M2	-0.20	1.64
Upper key	15		M3	-0.20	2.39

Predicted key	Stim. Num	Melodic sequence	Ending type	Mean Score	Std. dev.
Ambiguous	6		M3	0.20	1.48
Ambiguous	8		M3	0.20	1.48
Ambiguous	10		M3	0.40	2.07
Ambiguous	21		M2	0.40	1.52
Ambiguous	2		M3	0.60	2.51
Upper key	25		M2	0.60	3.85
Upper key	14		M2	0.80	3.49
Ambiguous	24		M2	1.00	1.22
Upper key	28		M2	1.00	2.83
Ambiguous	29		M2	1.20	2.17
Upper key	5		M3	2.00	2.83
Upper key	9		M3	2.00	2.92

Predicted key	Stim. Num	Melodic sequence	Ending type	Mean Score	Std. dev.
Upper key	17		M2	2.60	3.13
Upper key	19		M3	3.00	0.71
Upper key	1		M3	4.00	0.00

Note. The melodic sequences are displayed in the upper key of G major and lower key of C major, though actual materials were transposed across keys. M2 = major second, M3 = major third ending type.

Figure 1.

Left: The five sequences that most evoked the lower key. Right: The five sequences that most evoked the upper key. Sequences shown here are transposed to the pitch set [C, D, E, F, F#, G, A, B].

Figure 2.

Top: Estimated timescales for music and speech processing. Note that “mean syllabic rate” corresponds acoustically to the peak of the modulation spectrum for speech. Bottom: Average percent correct for each tempo in blue and average d' for each tempo in red. Error bars indicate estimated standard error.

Table 1

Results of Tukey-Kramer post-hoc comparisons for Experiment 2.

Level	Tempo (BPM)	Rate (Hz)	Inter-Onset Interval (ms)	Significant Comparisons
1	7	0.1	8571	5–9, 16–17
2	15	0.3	4000	5–9, 16–17
3	30	0.5	2000	12–17
4	45	0.8	1333	12–17
5	60	1.0	1000	1–2, 11–17
6	75	1.3	800	1–2, 12–17
7	95	1.6	632	1–2, 12–17
8	120	2.0	500	1–2, 11–17
9	200	3.3	300	1–2, 11–17
10	400	6.7	150	12–17
11	600	10.0	100	5, 8–9, 16–17
12	800	13.3	75	3–10, 16–17
13	1000	16.7	60	3–10, 17
14	1200	20.0	50	3–10, 17
15	1600	26.7	38	3–10, 17
16	2200	36.7	27	1–12
17	3400	56.7	18	1–15