Space Oddity: Musical Syntax is Mapped onto Visual Space

Neta Batya Maimon (netacoh3@mail.tau.ac.il)
Tel Aviv University
Dominique Lamy
Tel Aviv University
Zohar Eitan
Tel Aviv University

Research Article

Keywords: Tonality, Cross-modal correspondence, metaphor, cross-domain mapping, Implicit Association Test

Posted Date: May 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-528833/v1

Version of Record: A version of this preprint was published at Scientific Reports on November 16th, 2021. See the published version at https://doi.org/10.1038/s41598-021-01393-1.

Space Oddity: Musical Syntax is Mapped onto Visual Space

Neta B Maimon[1], Dominique Lamy[1,2] & Zohar Eitan[3]

[1]*The School of Psychological Sciences, Tel Aviv University, Tel Aviv, 69978, Israel*

[2]*The Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, 69978, Israel*

[3]*The Buchmann Mehta School of Music, Tel Aviv University, Tel Aviv, 69978, Israel*

* *corresponding author: netacoh3@mail.tau.ac.il*

ABSTRACT

Musicians ubiquitously apply spatial metaphors when describing the stability hierarchy established by tonal syntax: stable tones are considered spatially central and, as gravitational foci, spatially lower. We investigated whether listeners, musicians and non-musicians, indeed associate tonal relationships with visuospatial dimensions, including spatial height, centrality, laterality, and size, and whether such mappings are consistent with tonal discourse. We examined explicit and implicit associations. In the explicit paradigm, participants heard a tonality-establishing prime followed by a probe tone and coupled each probe with a subjectively appropriate location on a two-dimensional grid (Exp. 1) or with one of 7 circles differing in size (Exp. 4). The implicit paradigm used a version of the Implicit Association Test to examine associations of tonal stability with vertical position (Exp. 2), lateral position (Exp. 3) and object size (Exp. 5). Tonal stability was indeed as- sociated with perceived physical space: the spatial distances between the locations associated with different scale-degrees significantly correlated with the tonal stability differences between these scale degrees. However, inconsistently with the hypotheses implied by musical discourse, stable tones were associated with leftward and higher spatial positions, relative to unstable tones, rather than with central and lower spatial positions. We speculate that these mappings are influenced by emotion, embodying the "good is up" metaphor, and by the spatial structure of music keyboards. Taken together, results suggest that abstract syntactical relationships may consistently map onto concrete perceptual dimensions across modalities, demonstrating a new type of cross-modal cor- respondence and a hitherto under-researched connotative function of musical structure.

Keywords: Tonality, Cross-modal correspondence, metaphor, cross-domain mapping, Implicit Association Test

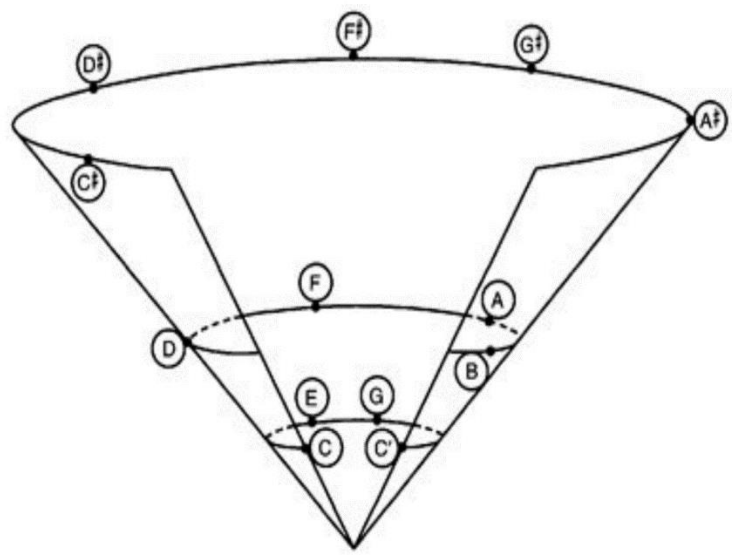
Introduction

Since the 17th century, the structure of most Western music has been strongly governed by tonality. Tonality is a system of interrelationships among tones and chords within major or minor keys, all relating to a single tone and chord, the tonic, that is endowed with maximum stability and closure (for a brief primer on tonality, see Appendix 1). In important ways, tonality is a musical syntax: a set of rules and practices, implicitly understood by listeners, which governs closure and continuity, stability and tension, in melodic and harmonic sequences [1, 2]. It thus serves to shape listeners' melodic and harmonic expectancies, and governs their evaluation of how well a given tone or chord fits a musical context, as well as their evaluation of the musical tension such tone evokes (for reviews of tonal cognition research see [3-5])

Notably, tonality is not merely an abstract cognitive schema: it plays an important role in establishing music's emotional connotations. For instance, unstable tonal progressions, which imply continuation, tend to evoke higher arousal and more negative valence, compared to stable, closural progressions (e.g., [6-8]) A recent study by our group [9] suggests that tonality's connotations may extend beyond emotional dimensions, and map onto aspects of the perceived physical world. We showed that listeners systematically associate more stable, closural tonal progressions with brighter visual stimuli, both explicitly and implicitly. We thus demonstrated that tonal syntax may engender systematic cross-modal associations.

Such associations are particularly intriguing. While cross-modal associations of basic auditory parameters, like pitch and loudness, have been widely investigated and corroborated (for research reviews see [10-12]), cross-modal associations of tonal syntax, as shown by Maimon et al. [9], have not - and they suggest a different, novel domain of cross-modal research. Such relationships imply that an abstract syntactic structure may systematically map onto perceptual dimensions, and thus that syntactic relationships in music may also carry concrete semantic connotations. In the present study, we further explore this fascinating possibility, by examining how tonal syntax is associated to perceived physical space.

Spatial analogies have often served to model and illustrate tonal structures and relationships. For instance, Krumhansl [13] depicts the stability hierarchy and perceived similarities among melodic scale degrees in a spatial map (Figure 1). Obviously, such renditions of "tonal pitch space" [14] were not meant to be interpreted literally, in perceptual terms: Figure 1, for example, does not suggest that the tone C# is actually perceived as located above and to the left of F. Yet, it is such literal sense of perceived tonal space that the present study explores. We examine whether tonal relationships are actually associated by listeners (both musicians and non-musicians) with physical space, visually perceived: for instance, are more stable scale degrees perceived as situated above or below, to the right or to the left, of instable scale degrees?



An important motivation for such examination is the fact that musicians habitually conceptualize tonal relationships via spatial and spatio-kinetic metaphors (similarly to metaphorical mappings of other abstract domains in terms of physical relationships; e.g., [15, 16]). In fact, one can hardly refer to tonal relationships without recourse to terms derived from the spatial domain, such as "tonal center" (for the tonic note or chord) "leading tone" (for the note just below the tonic, commonly regarded as "pulling" toward it), or "cadence" (from the Latin cadere, to fall), a term for a closural tonal progressions. Similarly, tonal relationships were discussed and modelled by music theorists in terms of forces operating in physical space such as gravity or magnetism, or via basic spatial relations such as centre/periphery or top/bottom (for a historical review, see [17]).

Such spatial discourse on tonality raises specific hypotheses regarding the associations of tonal stability and spatial dimensions. Two prominent (and often linked) metaphors for tonal stability, both historically and in musicians' discourse, are the center/periphery mapping, where stable scale degrees, primarily the tonic (”tonal center”) are more central than unstable degrees (e.g., [13, 18, 19]), and the gravity metaphor, in which the tonic and other relatively stable tones or chords generate a “gravitational” pull on unstable tones or chords (e.g., [19, 20, 21]; For overviews of the use of metaphor in music discourse, see [22, 23]). Note that the gravity metaphor implies that stable tones are localized below the unstable tones they pull, an implication reflected in the accepted association of closure with physical descent by musicians (as in the cadence etymology; see also [24]). These mappings suggest two hypotheses associating tonality and perceived spatial location: stable tones should be more central or lower in space than unstable tones.

Importantly, cross-modal perception has been shown to correlate with established cross-modal metaphors. For instance, the prevalent “light is good/dark is bad” metaphor biases brightness perception of positive and negative words [25]. Such effects have been widely demonstrated with regard to auditory dimensions: we actually perceive “high” pitches as higher in space, and sounds with louder “volume” as more voluminous (see ref. [10,11] for reviews of empirical research). Though the causal direction associating perception and metaphor is not always clear, effects of using specific language metaphors on perception have been reported (e.g., high/low vs. thin/thick for pitch; [26]). The ubiquitous use of spatial metaphors to describe and model tonality thus serves to motivate this study: as in other domains, such metaphors may have perceptual counterparts, which they reflect or engender.

The present study examines the associations of several aspects of perceived two-dimensional space with the stability hierarchy of melodic scale-degrees. In that hierarchy, whose psychological reality has been repeatedly established (see [3-5, 27] for reviews), the tonic note (scale degree 1) and, to a lesser extent, the other members of the tonic chord (scale degrees 3 and 5) are considered most stable and closural, suggesting a relatively weak tendency to proceed to another tone. The other diatonic scale degrees (2, 4, 6 and 7) are less stable, and imply continuation to their more stable neighbours, 1, 3, or 5; the remaining five chromatic ("out of key") tones are the least stable, and strongly imply continuation to adjacent diatonic (within-key) notes.

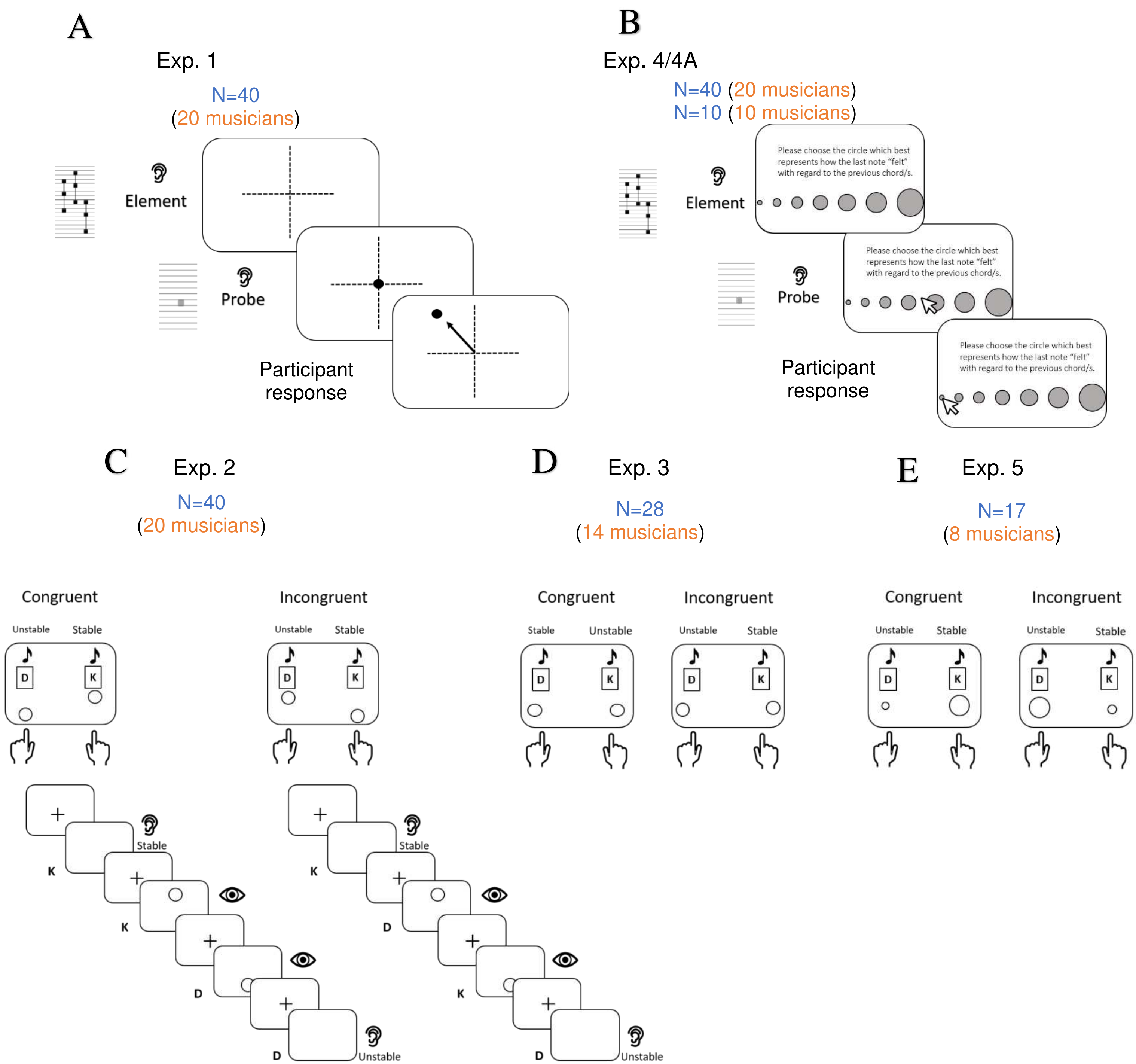
Two main research questions lead our examination of perceived tonal space. The first is whether, generally, tonal stability significantly associates with perceived space. If it is, then perceived spatial distances from the tonic note should correlate with tonal stability, such that more stable scale degrees (e.g., scale degrees 3 and 5) should be perceived as nearer to the tonic than unstable degrees, such as chromatic notes (Hypothesis #1).

A second, more specific question is whether correspondences of tonal stability and perceived space actually reflect mappings suggested by musicians' discourse -- particularly, tonal centricity and tonal gravitation. If they do, then stable tones should be perceived as located lower in physical space (Hypothesis #2) or in more central positions (Hypothesis #3) than less stable tones.

In addition to these dimensions, we also examined associations of tonal stability with lateral position (right/left), and physical size, as both dimensions were closely associated with auditory features in previous studies [10,11,12]. We have no priori hypotheses for these dimensions, however (although size, which often correlates with weight, could be associated with stability).

We investigated these variables in six behavioural experiments, using both explicit and implicit measures (see figure S3 for experimental design). The explicit paradigm (Experiments 1, 4, 4a) applied a version of the probe-tone method [28]. On each trial, participants heard a tonality-establishing prime (a chord sequence), followed by a probe tone (one of the 12 chromatic tones). In Experiment 1, we examined the association of tonal stability with vertical position (Hypothesis 1), lateral position, and centrality (Hypothesis 2). Participants were asked to associate each probe tone with the location they felt was most appropriate on a two-dimensional grid. In Experiments 4 and 4a, we examined the association of tonal stability with physical size. Participants matched each probe with one of 7 circles differing in size.

The implicit paradigm was a variant of Parise and Spence's [29] cross-modal version of the Implicit Association Test (IAT), that examined associations of tonal stability with vertical position (Experiment 2), lateral position (Experiment 3) and object size (Experiment 5). In these experiments, a tone (either tonally stable or unstable) and a visual image (e.g., a circle either high or low in space) were assigned to the same key on the computer keyboard in hypothetically congruent (e.g., stable/low, unstable/high) or incongruent (e.g., stable/high, unstable/low) combinations. Participants were asked to press the assigned key as fast as possible in response to each tone or image. Finding that performance is better for congruent than for incongruent combinations would provide evidence that tonal stability is associated with the spatial dimensions examined. To explore the role of conceptual musical knowledge on these associations, all experiments tested both musically trained and musically untrained participants.


A
Exp. 1
N=40
(20 musicians)
Element
Probe
Participant response
B
Exp. 4/4A
N=40 (20 musicians)
N=10 (10 musicians)
Please choose the circle which best represents how the last note "felt" with regard to the previous chord/s.
Element
Probe
Participant response
C
Exp. 2
N=40
(20 musicians)
D
Exp. 3
N=28
(14 musicians)
E
Exp. 5
N=17
(8 musicians)
Congruent
Incongruent
Unstable
Stable
D
K
Stable
Unstable

Results

Localization in two-dimensional space

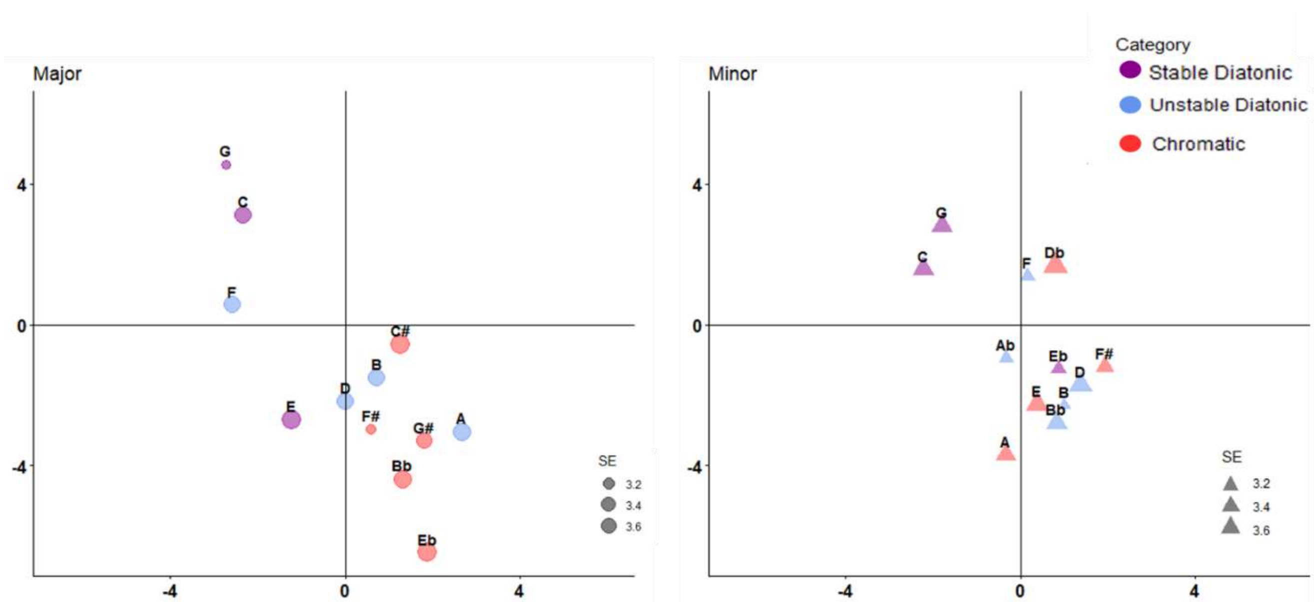
Experiment 1

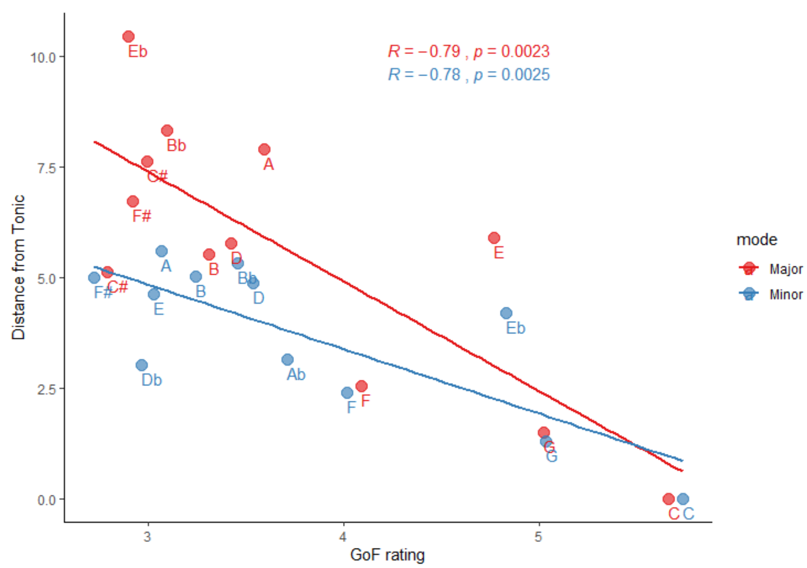
Forty participants (20 musicians) chose the subjectively appropriate location of each of the 12 chromatic scale degrees on a two-dimensional space, using a variant of the probe-tone method (28).

Mean localizations for each scale degree in Exp. 1 are presented in figure 2. We first addressed Hypothesis #1 by examining the correlation of the perceived spatial distances between scale-degrees and their tonal stability. To do so, we first calculated the distance between the locations associated with each scale degree (12 in major and 12 in minor) and the location associated with the tonic (the maximally stable degree). We then correlated these distances with mean goodness of fit (GoF) ratings, independently obtained in Maimon et al. (ref [9], Exp. 1), which used identical auditory stimuli. These ratings, which are similar to those obtained in "classic" GoF experiments [e.g., 28], closely correspond with the stability hierarchy commonly suggested by tonal theories. The obtained Pearson correlations validated our hypothesis: scale degrees with lower GOF ratings (i.e., which were rated as less fitting to a tonal context), were also located further away from the tonic: $r=-0.79$, $p=0.002$, and $r=-0.78$, $p=0.0025$ for major and minor modes, respectively (see figure 3).

Next, to explore the centrality hypothesis (Hypothesis #3), we calculated the radial distance of each scale degree from the axes' origin (0,0 point) and averaged it across the three tonal stability categories (i.e., stable diatonic, unstable diatonic and chromatic, for further details see materials and methods). Participants located the second tonal stability category-- the unstable diatonic tones -- closest to axes origin ($R=1.46$), stable diatonic tones were located further ($R=2.07$), and the chromatic tones were the farthest ($R=2.78$). These results thus invalidate the centrality hypothesis, since the most stable tones were not located closest to axes origin.

Finally, we separately examined how tonal stability associates with vertical and horizontal orientations. Participants' responses were analyzed with two mixed linear models, one for the localizations on the x axis and one for localizations on the y axis. The mixed linear models' variables of Exp. 1-3 are presented in table 1. In contrast to the Hypothesis #2, according to which stable scale degrees should be located lower in space, participants chose significantly higher spatial positions for more stable scale degrees (stable diatonic vs. unstable: t=-3.15, p=.002, stable diatonic vs. chromatic: t=-2.75, p=.008). Participants also localized more stable scale degrees in more left-hand positions (stable diatonic vs. unstable: t=2.84, p=.007, stable diatonic vs. chromatic: t=3.2, p=.002).

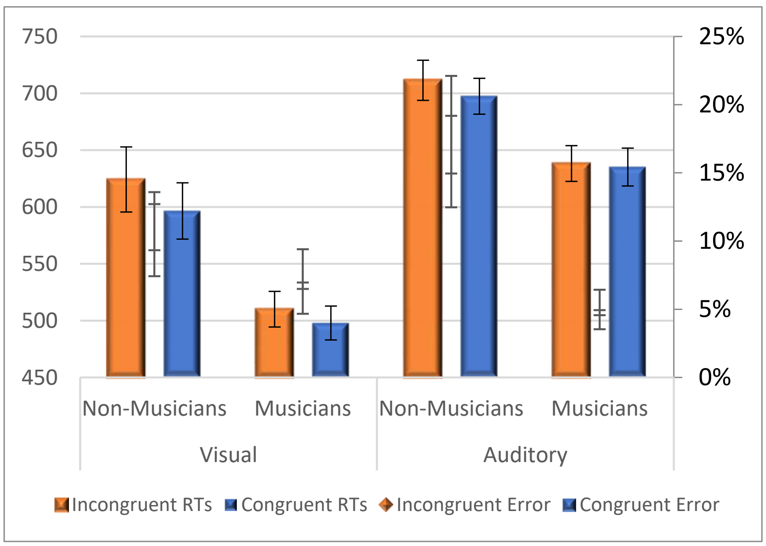


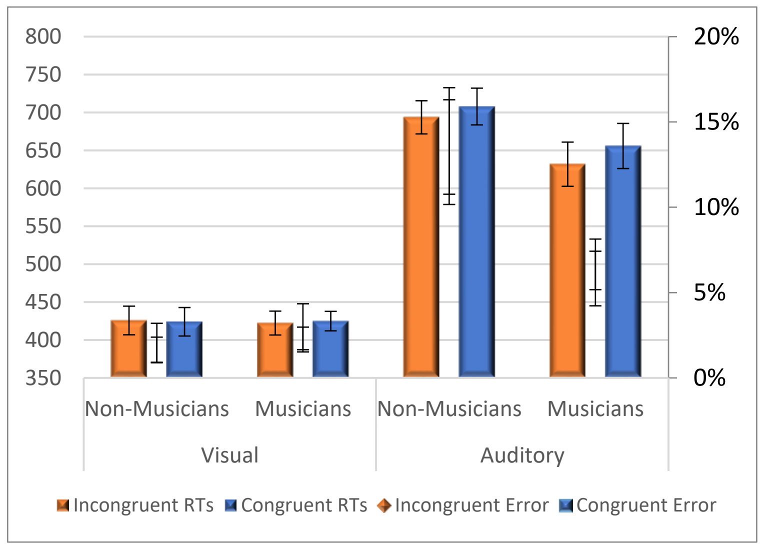


Experiments 2 & 3

Participants who underwent Exp 1 also participated in the two cross-modal versions of the Implicit Association Test (IAT), [29]. These two IAT experiments were designed to measure the implicit associations between tonal stability (stable/unstable auditory stimuli) and vertical height (high/low circle location, 40 participants, Exp. 2), or laterality (right-hand/left-hand positions, 28 participants, Exp. 3). The congruency conditions in these IATs were specified following the results obtained on Exp. 1: In Exp. 2, blocks where reaction keys to stable auditory stimuli were paired with higher circle location, and unstable stimuli with lower location were congruent. Likewise, in Exp. 3, blocks where reaction keys to stable auditory stimuli were paired with left-hand position, and unstable stimuli with right-hand position were congruent. Mean reaction times and error rates for the congruent and incongruent conditions of Exps. 2 and 3 are presented in figure 5. For each IAT experiment, reaction times were analyzed with mixed linear models and accuracy rates with generalized mixed linear models (see table 1 for mixed linear models results).

On the implicit test probing the vertical axis (Exp. 2), participants were faster (t=2.53, p=.015) and more accurate (z=-5.22, p<.001) in congruent blocks (i.e., where the response to the stable auditory stimuli was matched to higher circle location and the unstable stimuli to the lower location) than incongruent blocks. This effect was similar for musicians and non-musicians and for responses to both visual and auditory stimuli. The results for the implicit test probing the horizontal axis (Exp. 3), showed an opposite association relative to the explicit test (Exp. 1): participants were faster on incongruent blocks (i.e., when the response to the stable auditory stimuli was matched to the right-hand position and the unstable stimuli to the left-hand positions), than on congruent blocks ($t=4.11$, $p<.001$). However, this effect was less consistent, as it was modulated by modality ($t=-3.68$, $p<.001$), it was significant in responses to auditory stimuli ($t=-3.0$, $p=.02$) but not in responses to visual stimuli ($t<1$) and did not occur on accuracy rates (all ps>.05).





Experiment	Dependent Variable	Fixed Variable	Estimate	Std. Error	df	t value/ z value	Pr(>|t|)/ Pr(>|z|)
Exp 1	Y axis	(Intercept)	52.43	1.72	39.94	30.43	<0.001
		Category: unstable diatonic	-3.60	1.14	43.21	-3.15	**0.003**
		Category: chromatic	-3.42	1.24	38.58	-2.75	**0.009**
		Group: non-musician	-2.15	2.44	40.18	-0.88	0.384
		Category: unstable diatonic X Group: non-musicians	2.57	1.63	44.61	1.58	0.121
		Category: chromatic X Group: non-musicians	-0.81	1.76	39.48	-0.46	0.649
	X axis	(Intercept)	48.02	0.92	39.61	52.14	<0.001
		Category: unstable diatonic	2.28	0.80	39.18	2.84	**0.007**
		Category: chromatic	3.43	1.07	39.68	3.20	**0.003**
Exp 2	RTs	(Intercept)	632.94	15.59	43.37	40.59	<0.001
		Congruence: incongruent	13.28	5.24	39.73	2.53	**0.015**
		Modality: visual	-114.90	15.84	39.72	-7.26	**<0.001**
		Group: non-musicians	71.63	20.99	39.99	3.41	**0.001**
	Accuracy	(Intercept)	3.33	0.28		11.96	<0.001
		Congruence: incongruent	-0.24	0.05		-5.22	**<0.001**
		Modality: visual	-0.37	0.27		-1.36	0.173
		Group: non-musicians	-1.09	0.39		-2.83	**0.005**
		Modality: visual X Group: non-musicians	0.76	0.37		2.06	**0.040**
Exp 3	RTs	(Intercept)	663.61	19.29	28.21	34.40	<0.001
		Congruence: congruent	20.83	5.07	48.92	4.11	**<0.001**
		Modality: visual	-238.63	18.34	28.90	-13.01	**<0.001**
		Congruence: congruent X Modality: visual	-17.98	4.89	14577.37	-3.68	**<0.001**
	Accuracy	(Intercept)	3.05	0.28		10.74	<0.001
		Congruence: congruent	-0.05	0.07		-0.64	0.523
		Modality: visual	1.34	0.30		4.42	**<0.001**
		Group: non-musicians	-0.77	0.39		-1.96	**0.050**
		Modality: visual X Group: non-musicians	1.45	0.43		3.34	**<0.001**

Physical size

Experiments 4 & 4A

Forty participants (20 musicians) who did not participate in Exps. 1-3, matched each of the chromatic scale degree to a circle's size using a variant of the probe-tone method [28]. Each participant was presented with a row of seven circles ranging from smallest to largest, either from right-to-left or left-to-right (manipulated between-participants).

The results showed that all main effects and interactions were significant (see table 1 for descriptive statistics, and table S2 for mixed-linear models results). However, when breaking the significant interactions into different LMM models, the results showed that only musicians who were presented with the large circle on the left matched larger circles to more stable the tones, ($t=4.20$, $p=.001$, and $t=4.75$, $p<.001$, for stable diatonic vs. unstable diatonic, and unstable diatonic vs. chromatic, respectively). The remaining participants (i.e., musicians who were presented with the large circle on the right as well as both groups of non-musicians) showed no consistent association bias between size and tonal stability. Before we try to explain this discrepancy, we sought to replicate this finding.

Since the critical effect was found only in musicians, who possess conceptual musical knowledge, we conducted a follow-up experiment (Exp. 4A) on 10 additional musicians. Unlike in the main experiment, the direction of the circles' presentation was manipulated within participant. Each musician underwent two experimental sessions, one with larger circles presented to the left of smaller circles, and the other with larger circles presented to the right of smaller circles (and which came first was counterbalanced). The main LMM model results showed that larger circles were chosen for more stable tones across sessions. However, this effect was modulated by the direction of the circles' presentation. Simple effects analyses revealed that the stable diatonic tones were matched with larger circles than unstable diatonic tones only in sessions where larger circles were presented to the left of smaller circles, ($t(1,35)=3.21$, corrected $p=.042$). The comparison between unstable diatonic and chromatic tones on the left-to-right presentation, and all tonal stability effects when the same musicians were presented with the circles from right to left, did not reach significance (all $ps>.05$). Thus, we replicated the finding that musicians associate larger circles to more stable tones only when larger circles are presented on the left.

Experiment	Group	Circles Direction	Stable Diatonic	Unstable Diatonic	Chromatic
4	Musicians	Left --- Right	**5.43** ±1.74	**3.71** ±1.76	**3.46** ±1.73
		Left --- Right	**4.46** ±1.75	**4.27** ±1.7	**4.27** ±1.95
	Non-musicians	Left --- Right	**3.96** ±1.7	**3.91** ±1.75	**3.9** ±1.68
		Left --- Right	**4.07** ±1.73	**4.02** ±1.77	**3.96** ±1.77
4A	Musicians	Left --- Right	**4.58** ±1.6	**3.98** ±1.72	**4.03** ±1.67
		Left --- Right	**4.44** ±1.51	**4.33** ±1.67	**4.22** ±1.63

Experiment 5

Seventeen participants (8 musicians) underwent an IAT experiment associating circle size and tonal stability (Exp. 5). Participants were equally fast and accurate, irrespective of whether the response to the stable auditory stimuli was matched to the larger or smaller circle. Taken together with the finding that only musicians associated physical size and tonal stability (Exp. 4), the absence of any such association in the implicit test suggests that the association of size and tonal stability depends on explicit, conceptual musical knowledge.

Experiment	Dependent variable	Fixed variable	Estimate	Std. Error	df	t value/ z value*	Pr(>|t|)/ Pr(>|z|)
Exp 4	Size matching	(Intercept)	5.43	0.24	40.00	22.84	<.001
		Category: unstable diatonic	-1.71	0.29	40.01	-5.88	**<.001**
		Category: chromatic	-1.97	0.36	40.01	-5.52	**<.001**
		Large circle position: right	-0.96	0.34	40.00	-2.87	**0.007**
		Group: non-musicians	-1.46	0.34	40.00	-4.36	**<.001**
		Category: unstable diatonic X Large circle position: right	1.52	0.41	40.01	3.68	**0.001**
		Category: chromatic X Large circle position: right	1.78	0.50	40.01	3.53	**0.001**
		Category: unstable diatonic X Group: non-musicians	1.66	0.41	40.01	4.03	**<.001**
		Category: chromatic X Group: non-musicians	1.90	0.50	40.01	3.78	**0.001**
		Large circle position: right X Group: non-musicians	1.07	0.48	40.00	2.25	**0.030**
		Category: unstable diatonic X Large circle position: right X Group: non-musicians	-1.51	0.58	40.01	-2.60	**0.013**
		Category: chromatic X Large circle position: right X Group: non-musicians	-1.82	0.71	40.01	-2.55	**0.015**
Exp 4A	Size matching	(Intercept)	4.59	0.16	12.11	27.99	<.001
		Category: unstable diatonic	-0.65	0.14	15.86	-4.65	**<.001**
		Category: chromatic	-0.54	0.16	13.81	-3.35	**0.005**
		Large circle position: right	-0.14	0.11	4164.51	-1.31	0.189
		Category: unstable diatonic X Large circle position: right	0.54	0.14	3918.59	3.84	**<.001**
		Category: chromatic X Large circle position: right	0.32	0.14	4130.42	2.27	**0.023**
Exp 5	RTs	(Intercept)	521.53	10.55	12.86	49.43	<.001
		Congruence: congruent	3.66	6.01	15.82	0.61	0.551
		Modality: visual	-38.60	6.88	8.11	-5.61	**<.001**
	Accuracy	(Intercept)	2.47	0.37		6.73	<.001
		Congruence: congruent	-0.24	0.18		-1.36	0.174
		Group: non-musicians	-0.48	0.50		-0.96	0.336
		Modality: visual	-0.16	0.23		-0.70	0.487
		Congruence: congruent X Group: non-musicians	0.16	0.24		0.66	0.511
		Congruence: congruent X Modality: visual	-0.04	0.18		-0.22	0.827
		Group: non-musicians X Modality: visual	0.97	0.32		3.05	**0.002**
		Congruence: congruent X Group: non-musicians X Modality: visual	-0.26	0.25		-1.07	0.286

Discussion

Two main findings of this study stand out. First, participants -- musicians as well as non-musicians -- systematically associated aspects of physical space with tonal stability. Stability differences between scale degrees strongly correlated with their physical distance from each other, as inferred from their perceived position; and tonally stable scale degrees were localized higher and to the left of unstable degrees. Second, the associations that emerged here between tonal stability and spatial dimensions substantially differed from those suggested by musicians' discourse, which led us to predict that tonally stable scale-degrees would be localized at lower (Hypothesis #2) or more central (Hypothesis #3) positions in 2-dimensional space, none of which was supported.

The finding that stable scale degrees were localized higher than instable ones was particularly robust. It occurred in musicians and non-musicians alike, in both explicit (Exp.1) and implicit (Exp. 2) tasks, suggesting that the association is not mediated by reflective conceptual processes. However, as noted, this finding sharply contrasts with our hypothesis (#2): Clearly, participants systematically intuited a relationship between tonal stability and vertical position; yet, just as clearly, that intuition had little to do with established spatial metaphors (tonality as gravity, cadences "falling", closural melodic lines descending towards the tonic). What, then, might be its source?

We speculate that this mapping can be explained, at least in part, via a conceptual metaphor of a considerably wider scope than the domain-specific metaphors for tonality: the "good is up" metaphor. Spatial height or spatial rise correlate with positive valence in language and cognition. This is suggested both by commonplace language metaphors or idioms associating the two domains (e.g., I am feeling down or high, cheer up, it's going downhill; refs 15, 30), and by behavioral measures. For instance, the valence of words presented to participants affects spatio-visual attention: positive words shift attention upwards, and negative words shift attention downwards [31]. Correspondingly, moving objects up or down enhances recall of positive and negative episodic memories [32-34] and learning of novel positive and negative words, respectively [35,36].

As noted, tonal stability has also been related to valence [6-8]. Particularly relevant to the present findings is a recent study by our group [37]. In that study, we applied explicit and implicit paradigms similar to those used here, with the same auditory stimuli, in order to examine associations of melodic scale degrees and emotional valence, instantiated by facial expressions. These associations emerged on both explicit and implicit tests: more stable scale degrees were associated with happier facial expressions. We may suggest, then, that the association we found here between tonal stability and vertical position may be mediated by emotional valence: tonally stable is good (i.e., emotionally positive); ergo, tonally stable is up. This interpretation is further supported by an increasing number of empirical studies suggesting that cross-modal correspondences may be mediated by emotion in musical contexts, such that a musical feature and a non-auditory feature (e.g., a specific color or smell) are related via shared association with the same emotion ([38-41]; Also see [9], for a comparable interpretation of the association of tonal stability with visual brightness).

A second correspondence emerging from this study involves lateral position: stable scale degrees were localized to the left of unstable degrees. This correspondence was observed for both musicians and non-musicians, though only in the explicit task (Exp. 1); unlike the vertical location association discussed above, it seems to stem from music-specific connotations. We suggest that the left-hand localization of tonal stability is mediated by pitch height, itself associated with both stability and lateral position. First, tones lower in pitch are associated with left-side spatial positions -- though mainly by musicians, since the association probably stems from the structure of musical keyboards (see [10] for research review). Second, in tonal music, the bass voice, lowest in pitch, is in several respects the most stable: bass notes tend to concur with stable metrical positions, coinciding with the metrical beat; they also tend to have a longer inter-onset duration than higher-voice notes, and are thus durationally stabler [42, 43]; and instable, "ornamental" melodic progressions tend to be featured in upper voices, rather than in the bass. Thus, the laterality association of tonally stable notes may be mediated by pitch height: stability is associated with low pitch, itself associated left-hand position.

This interpretation nicely dovetails the results of Exps. 4 & 4A, which examined the correspondence of tonal stability and size. Though size was not generally associated with tonal stability in these experiments, a highly significant interaction emerged: for musicians, stabler scale degrees were associated with larger objects – but only when larger objects were situated to the left. Notably, larger size has repeatedly been associated in cross-modal research with lower auditory pitch (see [10,11] for research reviews). Hence, the correspondence large/left-stable suggests in yet another way an association of left spatial position, lower ("larger") pitch, and tonal stability. Tonally stabler notes are larger, but only when associated with lower pitch height via its lateral position on musical keyboards.

Beyond specific correspondences, the most noteworthy finding of this study is its broadest: listeners map tonal hierarchy onto perceived physical space. These mappings, which do not seem to stem from established musical discourse, are mostly shared by musicians and non-musicians, suggesting that they are not dependent upon musical expertise or a conceptual music-theoretical knowledge. Strikingly, participants with no conceptual musical knowledge, asked to map tonal stimuli spatially, intuitively localized these stimuli such that their spatial distances closely correlated with their tonal stability differences. Furthermore, the most robust correspondence found here, between tonal stability and vertical position, emerged via an implicit paradigm (IAT) as well, possibly reflecting an automatic, involuntary association between the two dimensions.

The significance of these findings is threefold. For music, they suggest one path through which musical structure – often considered abstract and non-referential -- may refer to the perceived physical world, including its non-auditory dimensions. For cross-modal research they present (together with other findings of our group; ref. [9]) a novel type of correspondence: rather than associating basic perpetual dimensions in different sensory modalities (e.g., auditory loudness and visual brightness), they map abstract syntactical features onto concrete perceptual dimensions across modalities. Finally, our results demonstrate that syntactical structures may possess a connotative, semantic function – at least, but perhaps not solely, in the musical domain.

Materials and Methods

Ethical clearance was obtained for experimental data collection by Tel-Aviv University Institutional Review Board, and ethical guidelines followed.

Sample Size Selections: For experiments 1 and 4, we calculated the sample size required to observe a significant effect of tonal hierarchy, based on the study by Maimon et al. (ref. 9; Exp 1, GOF session). We based our power analysis on the comparison of GOF ratings between stable diatonic and chromatic tones. Using G*Power [44] with an alpha of 0.05, and a power of 0.80, we found the minimum required sample size to be 17 participants. For Exps. 2, 3 & 5 we calculated sample size based on Maimon et al. (ref. 9; Exp. 2, IAT of tonal stability and visual brightness). We found the minimum required sample size to be 28 participants.

Apparatus: The stimuli were presented on a computer with an Intel Core i7 CPU 920 processor, 2.67 GHz speed, and 2.98 GB RAM memory. Auditory stimuli were delivered through Sennheiser 210HD precision headphones. These were connected to the computer by a Terratec Producer, Phase 24 sound card. Stimulus loudness was measured through Brüel & Kjær type 2232 noise meter, measuring A-weighted dB. The computer screen was a 17-inch Lenovo LCD, with a screen refresh frequency of 85Hz and 1024 x 768-pixel resolution. The experiments were programmed and run using Matlab.

Experiment 1

Participants: Forty undergraduate and graduate students from Tel-Aviv University, 17 musicians (9 female, mean age = 20.93, SD = 2.74) and 20 non-musicians (12 female, mean age = 23.59, SD = 2.13) served as participants. The musicians had an average of 11.25 (SD = 3.4) years of musical experience (with a minimum of 7 years) and an average of 7.18 (SD = 2.09) years of music theory studies (with a minimum of 5 years). They all currently played and performed music. The non-musicians had an average of 2 (SD = 1.41) years of musical experience (with a maximum of 3 years in childhood), no music theory education, and none of them currently played music. All participants participated in the experiment for monetary credit

(25$ for two sessions of one hour, reported as Experiments 1 and 2).

Auditory stimuli: Pitch height is known to be strongly associated with visual space [45]. In order to minimize pitch-height effects, all auditory stimuli (both contexts and probes) were created using Shepard tones [46], Shepard tones are tones with a specific pitch chroma (pitch-class), sounded in 5 octaves simultaneously. Following the standard protocol presented in [46], we created the tones using a loudness envelope across the frequency range of 77.8-2349 Hz, with partials uniformly decreasing in intensity toward low and high frequencies. Shepard tones generate a clear pitch chroma (pitch class) but an ambiguous pitch height (i.e., the register in which a pitch is perceived – which spectral component in a given Shepard tone represents its pitch height for a listener -- is ambiguous and determined contextually). All stimuli were sounded in 71 dB.

Each trial consisted of a chord or a chord sequence which established a sense of tonal key -- the "context element" -- followed by a single tone out of the 12 chromatic tones in western tonality – the "probe". There were eight possible elements (see audio examples in supplementary materials 1): two element types (a triad, the first, third and fifth degrees played simultaneously, presumably perceived as a tonic chord, and a IV-V-I cadence, a sequence of three chords which imply a strong notion of closure in the tonal system). Each of these two types was played in four keys (G major, G minor, D-Flat major and D-Flat minor). There were 12 possible probes: the 12 tones of the chromatic scale (12 notes from C to B). On each trial, an element was played for .5 sec when it was a tonic chord, and for .5 sec for each chord followed by a silence of .25 sec when it was a cadence. Then, following a silence of 1 sec, a probe tone was played for .5 sec. Sound examples 1-4 present four examples of an element followed by a probe. In Sound example 1, an element (a cadence) in D-flat major is followed by a stable probe (the tonic note). In Sound Example 2 the same element is followed by an instable probe. In Sound Examples 3 and 4 elements in G major are followed by stable and instable probes, respectively.

The experiment consisted of 18 blocks. The first two blocks served as practice. The elements for these blocks were quasi-randomly drawn from the eight possible elements and consisted of either a cadence or a triad, in either G or D Flat (counterbalanced across subjects). They were followed by 16 experimental blocks, two for each of the eight possible elements, randomly mixed. The same element was used throughout any given block of trials. Each block consisted of 14 consecutive trials. The first two trials served as practice, using two randomly selected probes. In the following 12 experimental trials, the 12 chromatic tones, randomly mixed, were used as probe tones.

Procedure: The experiment was conducted in a dark sound-attenuating room. Participants were seated 50 cm from the computer screen. The experiment was divided into two sessions, the 2-dimensional matching-task session and the IAT task session, administered 7 to 14 days apart (M=8.62). Task order was counterbalanced across participants. Through the whole block, two dashed-line axes (x cm each), were presented at the middle of the screen. On each trial of Exp. 1, participants were instructed to listen to the element followed by the probe. Simultaneously with the auditory probe, a circle (2.5 cm diameter) appeared at the intersection of the 2 axes. Participants were asked to move the circle to the location that best matched the subjective judgment or "feeling," evoked by the probe tone (the circle was the cursor and participants moved it amd clicked at the selected location).

The experimenter underscored the subjective and intuitive character of the task. She instructed participants to choose whichever strategy they felt was the most suitable to perform the tasks, yet to be consistent and use the same strategy throughout the experiment. Additionally, participants were instructed to use the whole 2-dimensional space and to avoid placing the circle in the same quadrant in all trials.

After each block, participants were asked about their confidence in the judgments they had provided. Our objective was to examine whether mode and musical expertise (whether participants had professional musical training) affected how confident participants felt about their judgments. The text "How confident were you with the ratings of the last block?" appeared on the screen and participants provided a 1-7 rating by pressing the appropriate numeral on the upper row of the keyboard. Then, a 10-sec. white noise (71 db) was heard, followed by a count down from 5 to 1. Participants then proceeded to the next block by pressing any key, after a self-pace break.

All written and oral instructions were presented in Hebrew.

Experiment 2

Participants: Forty undergraduate and graduate students from Tel-Aviv University (20 musicians) who participated in Exp. 1. The order of the sessions was counterbalanced with a mean of X days between sessions (and a minimum of one week).

Auditory stimuli: The auditory stimuli in this experiment were designed to be comparable to the auditory stimuli in experiment 1 (i.e., element and probe, for auditory examples and graphical representations, see supplementary materials). Since responses in this experiment were speeded, we created a shorter and faster element consisting of a half-cadence (IV-V46-35) – a chord sequence suggesting a subsequent resolution by a stable chord or tone. Each chord of the element sequence lasted for 250 ms, with a 250mssilence apart. Since the IAT paradigm requires a speeded classification between two levels, we used the two probes that received the highest and lowest space ratings in Experiments 1A and 1B. Therefore, this element was followed by a 250 ms tone, either the tonic note (scale-degree 1, a tonally stable stimulus) or the raised subdominant (4#, a tonally unstable chromatic note). To minimize possible short-term memory effects, the cadence (played in a three-voice texture) was designed such that neither of the final probe tones (tonic, raised 4th) was included in the chords preceding it. The final tone was octave-doubled, and the pitch direction (up/down) between the last chord in the cadence and the following tone was controlled. Each stimulus was presented in two keys (C major, D-flat major), in a piano timbre (generated by Notion 6 music notation software). Sound Examples 2a and 2b in supplementary materials demonstrate stable and instable auditory stimuli used in this experiment.

Visual stimuli: The visual stimuli consisted of two circles, one with a presented 300 higher in the visual field and one 300 lower in the visual field.

Procedure: For trial sequence example, see figure 1C. On each trial, participants were presented with a unimodal stimulus (either auditory or visual). They were asked to classify it using one of two keys (K or D). For the visual stimuli, one key was assigned to the bright circle, and another to the dark circle. For the auditory stimuli, one key was assigned to a stable progression, and the other to an unstable progression. Note that the experimenter did not mention any of these adjectives (dark, bright, stable or unstable) but instead referred to the visual and auditory stimuli as type K and type D according to the participant's key assignment in the practice blocks. The experiment consisted of 24 blocks. All stimuli were presented in all blocks. Each block consisted of 28 consecutive trials. The first four trials served as practice trials and included one stimulus of each type (auditory stable, auditory unstable, visual bright and visual dark). In the following 24 experimental trials, the two visual stimuli were presented 6 times each, and the four auditory stimuli (stable and unstable in C and C# keys) were presented 3 times each, all randomly mixed. Each of the four possible stimulus-key pairings (2 congruent and 2 incongruent pairings) was presented in different blocks. There were 6 blocks for each pairing, resulting in 24 experimental blocks presented in randomly mixed order. Thus, in half of the blocks, the response pairing was congruent (i.e., the same response was associated with the stable auditory stimuli and the bright circle and with the unstable auditory stimuli and the dark circle), whereas in the remaining half of the blocks, the response pairing was incongruent (i.e., the same response was associated with the unstable auditory stimuli and the bright circle and with the stable auditory stimuli and the dark circle).

Experiment 3

Participants: 28 participants that underwent experiment 1 and 2 were invited for a third session. 14 musicians (7 females), mean age = 22.28, SD = 3.17 and 14 non-musicians (6 females), mean age = 23.78, SD = 1.52. The musicians had an average of 12.28 (SD = 2.71) years of musical experience (with a minimum of 7 years) and an average of 6.28 (SD = 1.57) years of music theory studies (with a minimum of 5 years). All participants participated in the experiment for monetary credit (15$).

Apparatus, stimuli, and procedure were the same as experiment 2. Visual stimuli in this experiment included circles were presented either 300 towards the left of the screen and 300 towards the right of the screen.

Experiment 4

Participants: Forty undergraduate and graduate students from Tel-Aviv University, 20 musicians (10 female) mean age = 22.82, SD = 3.21 and 20 non-musicians (15 female) mean age = 25.4, SD = 3.21 served as participants. The musicians had an average of 15.25 (SD = 3.15) years of musical experience (with a minimum of 7 years) and an average of 8.15 (SD = 3.02) years of music theory studies (with a minimum of 5 years). All participants participated in the experiment for monetary payment (15$).

Apparatus, stimuli, and procedure were the same as experiment 1. Visual stimuli in this experiment included 7 circles presented with varying size from small (x cm diameter) to large (x cm diameter). Half of the participants were presented with the small circle on the left and half with small circle on the right.

Experiment 4A

Participants: 10 musicians (3 females) from Tel-Aviv University who did not participate in Experiments 1-4, mean age of 24.41 SD=4.03, years of musical experience 12 SD=6.88, and 6.65 SD=3.47 years of music theory studies. All participants participated in the experiment for monetary payment (25$ for the two sessions)

Apparatus, stimuli, and procedure were the same as experiment 4. Procedure included two experimental sessions per participant, one with the small circle presented on the right and ont with the small circle presented on the left (order counter balanced).

Experiment 5

Participants: 17 undergraduate and graduate students from Tel-Aviv University, 8 musicians (6 females) mean age = 22.75, SD = 4.23 and 9 non-musicians (4 females) mean age = 23.88, SD = 3.92 served as participants. The musicians had an average of 11.75 (SD = 2.1) years of musical experience (with a minimum of 7 years) and an average of 5.87 (SD = 2.89) years of music theory studies (with a minimum of 4 years). All participants participated in the experiment for monetary credit (15$).

Apparatus, stimuli, and procedure were the same as experiment 2. Visual stimuli included a large circle (2 cm diameter) and small circle (0.5 cm diameter) presented at the middle of the screen.

Statistical analyses

Probe-tone experiments (Exp 1, 4 & 4A)

Independent variables included musical training (musician/non-musician) as a between-participants variable, and mode (major/minor) and tonal stability as within-participant variables. The tonal stability variable included three categories: Stable diatonic (the tonic chord members – the 1st, 3rd, and 5th degrees of the scale; that is, tones 1, 5, 8 of the 12 chromatic tones in major and 1, 4, 8 in minor), Unstable diatonic (2nd, 4th, 6th and 7th degrees of the scale: 3, 6, 10, 12 of the 12 chromatic tones in major and 3, 6, 7, 11, 12 in minor, where both options for the 7th scale-degree, raised [leading-tone] and natural, were included) and chromatic tones (non-scale degrees: 2, 4, 7, 9, 11 of the 12 tones in major and 2 ,5 ,7, 10 in minor). Dependent variables included mean position of X axis and mean positions of Y axis.

Two mixed-effect linear regression models (one for each dependent variable) with subjects as random effect were fit to the data in a step-wise-step up procedure. In these models, for variables with more than two levels, dummy variables are computed. Therefore, for tonal stability, which has 3 levels (stable diatonic, unstable diatonic and chromatic), we chose stable diatonic as the reference level, and two dummy variables were computed: unstable diatonic (e.g. the difference between stable diatonic and unstable diatonic), and chromatic (e.g. the difference between stable diatonic and chromatic). Accordingly, a significant effect of the unstable diatonic variable would mean that the level of the stable diatonic category was significantly different than that of unstable diatonic; similarly, a significant effect of the chromatic variable would mean that the level of the chromatic was significantly different than that of the stable diatonic. We used a step-wise-step up procedure on independent variables.

This procedure included incorporating the variables in importance descending order (i.e., from tonal stability to mode), and different forms of random slopes. All analyses were conducted via lme function [47] using R studio. For the final models of all mixed-linear models conducted in the present study see table 4.

IAT experiments (Exp 2, 3 & 5)

Independent variables included musical training (musician vs. non-musician) as a between-participants variable, and congruence (congruent/incongruent) and modality (visual/auditory) as within-participant variables. Dependent variables included reaction times (of correct trials), and accuracy rates. For reaction times (RTs), a mixed-effect linear regression models which contained the subjects as random effect were fit to the data in a ad-hoc step-wise-step up procedure (i.e., from congruency to modality, and different forms of random slopes. For final model see table 3).

For accuracy rates, a mixed-effect binomial logistic regression model which contained speaker as random effect was fit to the data in a step-wise-step up procedure. Fixed and random factors were inserted in the same order as for RTs model.

Experiment	Depended variable	Final model
1	X	X ~ Tonal Stability + (1+ Tonal Stability|Subject)
	Y	Y ~ Tonal Stability * Musical Experience + (1 + Tonal Stability | Subject)
2	RTs	RT ~ Congruence + Modality + Musical Experience + (1 + Congruence+Modality|Subject)
	Accuracy	Accuracy ~ Congruence + Modality * Musical Experience + (1+Modality|Subject)
3	RTs	RT ~ Congruence * Modality + (1 + Congruence+Modality|Subject)
	Accuracy	Accuracy ~ Congruence + Modality * Musical Experience + (1+Modality|Subject)
4	Size matchings	response ~ Tonal Stability * Circles Direction * Musical Experience + (1+ Tonal Stability|Subject)
4A	Size matchings	response ~ Tonal Stability * Circles Direction + (1+ Tonal Stability |Subject)
5	RTs	RT ~ Congruence + Modality + (1+Musical Experience+Modality+Congruence|Subject)
	Accuracy	Accuracy ~ Congruent * Musical Experience*Modality + (1+Modality+Congruence|Subject)

References

1. S. Koelsch, Toward a neural basis of music perception — a review and updated model, *Front. Psychol.* **2**, 110 (2011).
2. F. Lerdahl, R. Jackendoff, *A Generative Theory of Tonal Music*. (Cambridge: MIT, 1983).
3. C. L. Krumhansl, The cognition of tonality–as we know it today. *Journal of New Music Research*, **33**, 253-268 (2004).
4. M. A. Schmuckler, "Tonality and contour in melodic processing" In *Oxford library of psychology, The Oxford handbook of music psychology*. S. Hallam, I. Cross, & M. Thaut Eds. (Oxford University Press,2016), pp. 143–165.
5. D. Shanahan, "Musical Structure: Tonality, Melody, Harmonicity, and Counterpoint". In *The Routledge Companion to Music Cognition*. M. Cobussen, V. Meelberg, B. Truax, Eds. (Routledge, NY, NY, USA, 2016), pp. 141-151.
6. P. N. Juslin, Music and emotion: There is more to music than meets the ear. *Emotion Researcher* (2014).
7. S. Koelsch, T. Fritz, G. Schlaug, Amygdala activity can be modulated by unexpected chord functions during music listening. *Neuroreport*, **19**(18), 1815-1819. (2008).
8. N. Steinbeis, S. Koelsch, J. A. Sloboda, The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of cognitive neuroscience*, **18**(8), 1380-1393 (2006).
9. N. Maimon, D. Lamy, Z. Eitan, Crossmodal Correspondence Between Tonal Hierarchy and Visual Brightness: Associating Syntactic Structure and Perceptual Dimensions Across Modalities, *Multisensory Research*, ***33***(8), 805-836 (2020).
10. Z. Eitan, "How pitch and loudness shape musical space and motion" in The psychology of music in multimedia, S. L. Tan, A. J. Cohen, S. D. Lipscomb, R. A. Kendall, Eds. (Oxford university press. Oxford, UK, 2013), pp. 165-191.
11. Z. Eitan, "Cross-modal correspondences" in The Routledge Companion to Music Cognition, R. Ashley, R. Timmers, Eds. (Routledge, New York, NY, USA, 2017), pp. 213–224.
12. Spence, C. (2011). Crossmodal correspondences: A tutorial review, *Atten. Percept. Psychophys.* **73**, 971–995.
13. C. L. Krumhansl, The psychological representation of musical pitch in a tonal context. *Cogn. Psychol.*, **11**(3), 346-374 (1979).
14. F. Lerdahl, *Tonal Pitch Space* (Oxford University Press, New York, 2001).
15. G. Lakoff, M. Johnson, Metaphors we Live By (Chicago University Press, 1980).
16. E. Sweetser, *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure* (Vol. 54). (Cambridge University Press, 1990)
17. L .Rothfarb, "Energetics" *in The Cambridge History of Western Music Theory*, T. Christensen Ed. (Cambridge University Press, Cambridge, UK, 2001), pp. 927–955.
18. S. Larson, L. VanHandel, L. Measuring musical forces*, Music Percept*. 23, 119–136 (2005).
19. C. Brewer. A cognitive theory of musical meaning. *Journal of music theory*, **44**(2), 323-(2000).
20. J.P. Rameau. Traité de l'Harmonie Reduite à ses Principes Naturels: Divisé en Quatr Livres. (Jean-Baptiste-Christophe Ballard, Paris, France, 1722).
21. A. Schoenberg, Harmonielehre (Leipzig:Verlagseigentum der Universal-Edition, 1911); trans. as *Theory of Harmony*, by Roy E. Carter (Berkeley: University of California Press, 1978).

22. L. M. Zbikowski, *Conceptualizing Music: Cognitive Structure, Theory, and Analysis.* (Oxford University Press, 2002).

23. M. Spitzer, Metaphor and Musical Thought. (University of Chicago Press, 2004).

24. S. Larson. *Musical forces: Motion, metaphor, and meaning in music*. (Indiana University Press, Bloomington, 2012)

25. B. P. Meier, M. D. Robinson, L. E. Crawford, W. J. Ahlvers, When" light" and" dark" thoughts become light and dark responses: Affect biases brightness judgments. *Emo.*, **7**(2), 366 (2007).

26. S. Dolscheid, S. Shayan, A. Majid, D. Casasanto, The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psych. Sci.*, **24**(5), 613-621 (2013).

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

28. C. L. Krumhansl, E. J. Kessler, Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys, *Psychol. Rev.* **89**, 334–368 (1982).

29. C. V. Parise, C. Spence, Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test, *Exp. Brain Res.* **220**, 319–333 (2012).

30. G. Lakoff, M. Johnson, The metaphorical structure of the human conceptual system. *Cognitive science*, **4**(2), 195-208 (1980).

31. B. P.. Meier, M. D. Robinson, Why the sunny side is up: associations between affect and vertical position, *Psychological Science* 15: 243–247 (2004).

32. D. Casasanto, K. Dijkstra, Motor action and emotional memory. *Cog.*, **115**(1), 179-185 (2010).

33. S. Freddi, J. Cretenet, V. Dru, Vertical metaphor with motion and judgment: a valenced congruency effect with fluency. *Psych. Res.*, **78**(5), 736-748 (2014).

34. B. P. Meier, M. D. Robinson, The metaphorical representation of affect. *Metaphor and symbol*, **20**(4), 239-257 (2005).

35. D. Casasanto, A. de Bruin, Metaphors we learn by: Directed motor action improves word learning. *Cog.*, **182**, 177-183 (2019).

36. L. K. Globig, M. Hartmann, C. S. Martarelli, Vertical head movements influence memory performance for words with emotional content. *Front. Psych.*, **10**, 672 (2019).

37. N. Maimon, D. Lamy, Z. Eitan. Do Picardy thirds smile? Tonal hierarchy and tonal valence: explicit and implicit measures. arXive [preprint] (2021). https://arxiv.org/abs/2103.09709.

38. J. Bhattacharya, J. P. Lindsen, Music for a brighter world: Brightness judgment bias by musical emotion. *PloS one*, **11**(2), e0148959 (2016).

39. C. A. Levitan, S. Charney, K. B. Schloss, S. E. Palmer, The Smell of Jazz: Crossmodal Correspondences Between Music, Odor, and Emotion. In *CogSci* (2015).

40. S. E. Palmer, K. B. Schloss, Z. Xu, L. R. Prado-León, Music–color associations are mediated by emotion. *Proc. Natl. Acad. Sci. U.S.A.* **110**(22), 8836-8841 (2013).

41. K. L. Whiteford, K. B. Schloss, N. E. Helwig, S. E. Palmer, Color, music, and emotion: Bach to the blues. *i-Perception*, **9**(6), (2018).

42. Y. Broze, D. Huron, Is higher music faster? Pitch–speed relationships in Western compositions. *Music Perception*, **31**, 19-31 (2013).

43. H. Tamir & Z. Eitan. Higher is faster: Pitch register and tempo preferences. *Music Perception*, **33**, 179-198 (2015).

44. E. Erdfelder, F. Faul, F. A. Buchner, GPOWER: A general power analysis program. *Behavior research methods, instruments, & computers*, **28**, 1-11 (1996).

45. R. D. Melara, L. E. Marks, Interaction among auditory dimensions: Timbre, pitch, and loudness, *Percept. Psychophys*. **48**, 169–178 (1990).

46. R. N. Shepard, Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, **36**, 2346-2353 (1964).

47. Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Legends

Figure 1. Geometrical representation of perceived similarity between musical pitches within the octave in a tonal context. C major key serves as a reference, such that C represents the tonic (scale degree 1), C' is the pitch one octave above C. Taken from Krumhansl [13], Fig. 3.

Figure 2. Graphic representation of experimental designs in the present research. (A) Trial sequence in Experiment 1. (B) Trial sequence in Experiments 4 and 4A (C) Block opening screen and trial sequence of congruent (left) and incongruent (right) blocks in Exp. 2. (D) block opening screens of congruent (left) and incongruent (right) blocks in Exp. 3. (E) block opening screens of congruent (left) and incongruent (right) blocks in Exp. 5.

Figure 3. Mean localizations of each Major (left panel, circles) and Minor (right panel, triangles) scale degrees in semitones, in Exp. 1. Scale degrees are normalized into C Major and C minor (i.e., C = the tonic scale degree), # symbols represent raised degrees (e.g., F# == the triton), and b symbols represent lowered degrees (e.g., Db== lowered second), as a function of tonal stability category (stable diatonic =purple, unstable diatonic = blue and chromatic = red) and standard errors (relative size). Zoom-in to (-7,+7) on both axes.

Figure 4. Mean distance from the tonic of each scale degree in Major (red) and Minor (blue), as a function of GoF rating (independently obtained in ref. [9]), in Exp. 1. Scale degrees are normalized into C Major and C minor (i.e., C = the tonic scale degree), # symbols represent raised degrees (e.g., F# == the triton), and b symbols represent lowered degrees (e.g., Db== lowered second). Peasron R and *p* values correspond to mode (major/minor).

Figure 5. Mean reaction times (bars) and accuracy rates (diamonds) in Exp. 2 (left), cross-modal IAT between tonal stability and spatial height, and in Exp. 3 (right), cross-modal IAT between tonal stability and horizontal location, as a function of congruent (blue) and incongruent (red) blocks, by modality (visual, auditory), and musical experience (musicians, non-musicians). Congruency was determined according to Exp 1 results (e.g. tonality stable matched high space and left-hand positions). Results are presented with means and standard errors.

Table 1. Fixed effects results of Exp1, Exp2 and Exp3.

Table 2. Mean ± SD of size matchings in Exps. 4 & 4A as a function of musical experience group (musicians/non-musicians) and circles direction (large circle on the left/large circle on the right).

Table 3. Fixed effects results of Exps. 4, 4A and 5.

Table 4. list of the final mixed linear models in Exps. 1-5.

Acknowledgements

We thank Roni Granot, Liad Mudrik and Yeshayahu Shen for very useful discussions, Miriam Furst-Yust and the TAU School of Electrical Engineering for help with equipment, and Inbal Gur-Arie for her help in performing the experiments. This research was supported by an Israel Science Foundation grant (ISF 1920/16) to Zohar Eitan.

Author contributions statement

Z.E. and N.M conceived the experiments, N.M. conducted the experiments, D.L. and N.M. analyzed the results. All authors reviewed the manuscript.

Supplementary materials 1

Tonal Stability[1]

In music theory, the term 'tonality' denotes a system organizing pitch relationships in a musical work, both melodically (i.e., with regard to pitch succession) and harmonically (with regard to pitch simultaneities, or 'chords,' and their succession), in a hierarchy of stability and closure. Importantly, for each tonal context, or musical key, pitches are organized with reference to one maximally stable pitch class, the *tonic* or *tonal center*.

What is 'tonal stability'? Stable pitches and chords are associated with points of closure (endings of segments) in tonal music — an association underlying listeners' perception of musical closure [2,3]. Furthermore, unstable tones tend to proceed to the nearest stable tones (but not vice versa), a distributional fact affecting listeners' expectations [4]. Thus, unstable tones or chords tend to suggest to listeners a sense of tension and continuity, while the stable tones/chords following them suggest resolution of that tension [5]. Tonal stability also correlates with pitches' occurrence frequency: within a tonal context, stable pitches are more frequent than instable ones. Compatibly, listeners tend to perceive stable pitches as better fitting a primed tonal context than unstable ones [6].

In Western tonal music, tonal stability is governed by musical keys. A musical key is defined by its tonic and its mode (major or minor). The tonic — which may be any of the 12 pitch classes available in Western music — is, as noted above, the maximally stable, most closural note in a given key. For each tonic, two different modes are available, major or minor, each established by a different pattern of pitch intervals between the tonic and the other constituent tones. Thus, the keys of C major and C minor share the same tonic — the pitch class C — but have different modes; that is, the relationships of some of their other constituent pitches to the tonic differs.

Each major or minor key chiefly utilizes seven out of the 12 available pitch classes, its diatonic scale degrees. Diatonic scale degrees are commonly represented as a musical scale, arranged in order of pitch height, with the tonic note presented as first (lowest) note, or scale degree 1, and the other diatonic degrees presented and numbered accordingly (2–7; see Fig. S1A, B, for scalar representations in music notation of the two keys used in our experiments — G major and D-flat major). The most stable diatonic notes are the tonic (scale degree 1) and the other constituents of the tonic triad (the chord associated with the tonic note) — scale degrees 3 and 5. Other diatonic degrees (2, 4, 6, 7) are

[1] This section is taken from Maimon et al., (2020)

relatively instable, evoking tension that may be resolved when their more stable neighbors (1, 3, or 5 scale degrees) follow (5).

While simple tonal pieces (e.g., many western European folk songs) may exclusively use the seven diatonic scale degrees, more complex tonal music also applies the remaining five pitch classes (termed ‘chromatic tones’ by musicians), which are conceived (by music theorists) and perceived (by listeners; see ref 6, for review of empirical studies) as the least stable tones in a tonal context. These ‘out of key’ notes tend to evoke strong tension, resolved when their nearest diatonic (within-key) notes follow. In sum, then, musical theory, as well as music cognition research, distinguishes three main levels of tonal stability: tonal triad members (scale degrees 1, 3, 5) are the most stable; other diatonic notes (scale degrees 2, 4, 6, 7) are less stable, and tend to resolve to nearby stable scale degrees; and the remaining chromatic (out of key) notes are the least stable, strongly implying resolution by their diatonic neighbors. Figure S2 A demonstrates this hierarchy.

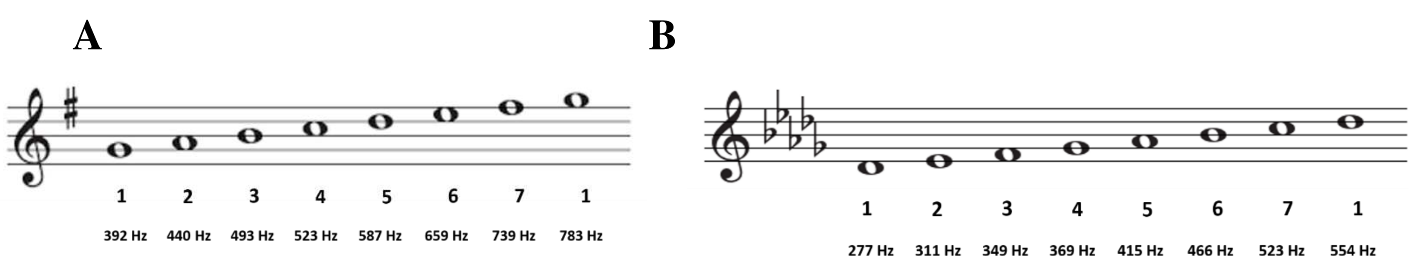


Figure S1. Musical notation of the two major scales used in Experiments 1A and 1B, G major (A), and D-flat major (B). Arabic numerals below the notes denote their scale degree identity (1–7). Fundamental frequencies (F0, in Hz) are marked below the scale degree numbers (note, however, that each scale degree can be represented by different pitch classes – notes sharing the same name, situated one or several octaves apart – e.g., A1; 55Hz; A2, 110Hz; A3, 220 Hz; A4, 440 Hz, etc).

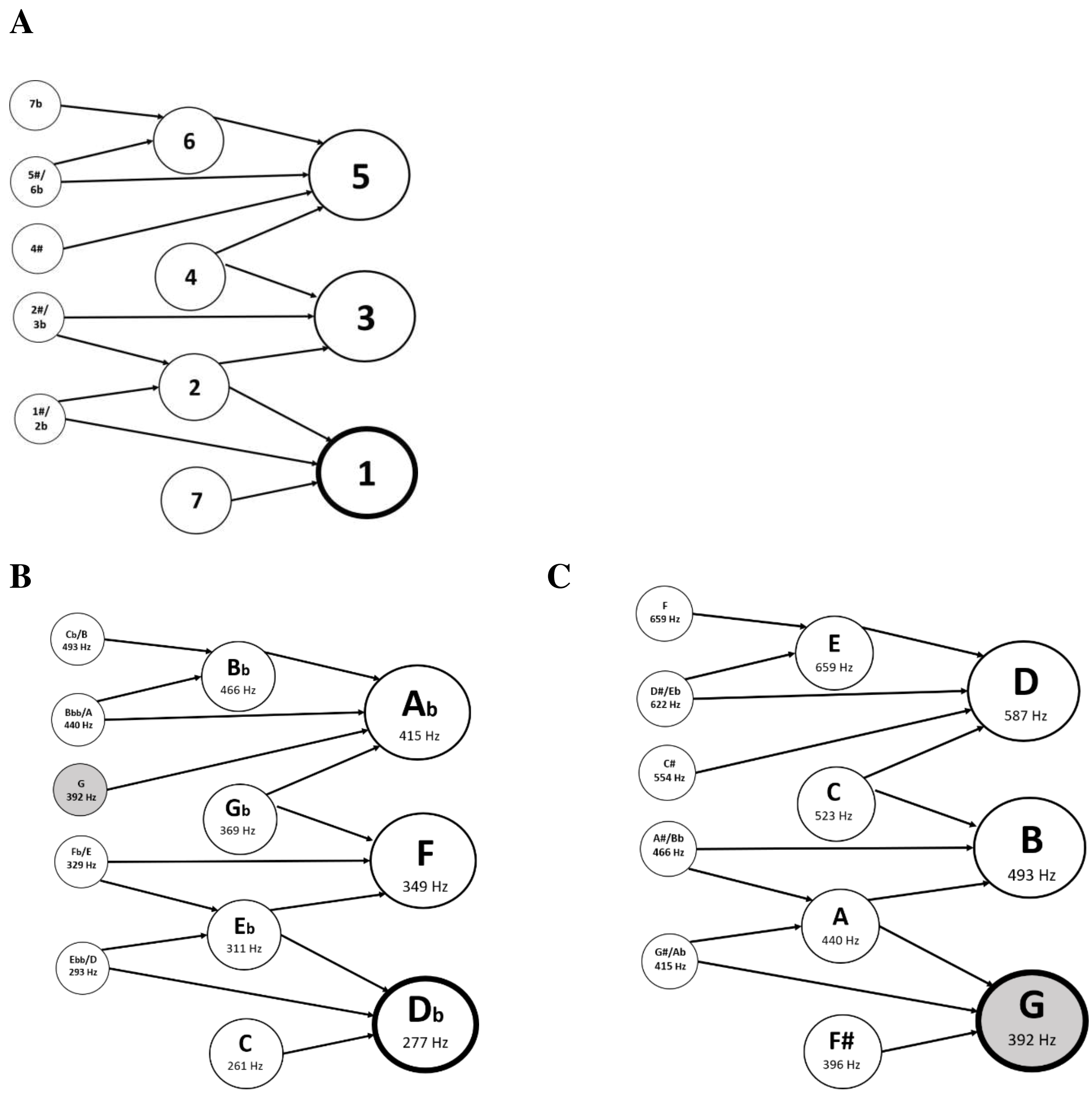


Figure S2. Visual representations of tonal hierarchy. Figure S2A presents generalized scale degree representation, while panels B and C, respectively, present these relationships in G major and D-flat major keys, replacing the scale degree numbers in A with the pitch classes representing these scale degrees in each key (fundamental frequencies are added below pitch class names, corresponding to those of pitches in Fig. S1). The larger and bolder scale degrees on the right are the most stable (stable diatonic), the middle layer presents other diatonic scale degrees — less stable yet belonging to the respective key (unstable diatonic), and the third (leftmost) layer presents the least stable tones (chromatic, ‘out of key’ tones). Arrows between scale degrees represent the typical motion direction between them — from an instable degree to a stable one. Note that while the tonal schema represented in 2A remains the same in both keys (2B, 2C), the representatives (i.e,. specific pitch classes) of its ‘slots’ (the scale degrees) radically change, with stable degrees in one key becoming instable in the other. For instance, the pitch class G (gray circles), which is the tonic note — the most stable note — in G major (2B) is an instable chromatic note in D-flat major (2C).

Importantly, the hierarchy of tonal stability is always relative to a key's tonic, such that the very same pitch may be highly stable in one key, and highly unstable in another. For instance, in the key of G major, the pitch class C-sharp/D-flat (which is a chromatic, 'out of key' note) would be extremely unstable, while G, the tonic, would be the maximally stable pitch class. In contrast, in the key of D-flat major, D-flat (the tonic note of that key) would be maximally stable, while G (here, a chromatic note) would be highly unstable. Figures 2A–C demonstrate this relativity: Fig. S2 A is a general representation of the tonal hierarchy described above. Figures S2 B and C demonstrate the same hierarchy as applied to two different keys — G major and D-flat major. Note that while the tonal schema remains the same in both keys, the representatives (i.e., specific pitch classes) of its 'slots' (the scale degrees) radically change, with stable degrees in one key becoming instable in the other.

Tonality underlies the bulk of Western music ('classical' as well as popular) since the 17th century and has been described and modeled in detail by music theorists (e.g., [7-9]). In recent decades, the psychological reality of tonality as a cognitive schema orienting the listener has been strongly established empirically (see 10-11 for research surveys). Studies applying converging experimental paradigms — explicit measurements, such as sung continuations and goodness-of-fit ratings (e.g., [12, 13 & 6]), as well as implicit ones, such as musical priming (e.g., 14) and event-related potentials (ERP; e.g., [15]) — have suggested that listeners implicitly abstract a tonal hierarchy, and the sets of melodic and harmonic expectancies it entails, closely matching those conjectured by music theorists' models.

Stimuli

Exp 1, 4 & 4A

Audio S1. Examples of stable (A-link) and unstable (B-link) auditory stimuli used in experiments 1, 4 and 4A. Both examples consist of a cadence type context element (a sequence of three chords), followed by either a stable (A) or an unstable (B) probe.

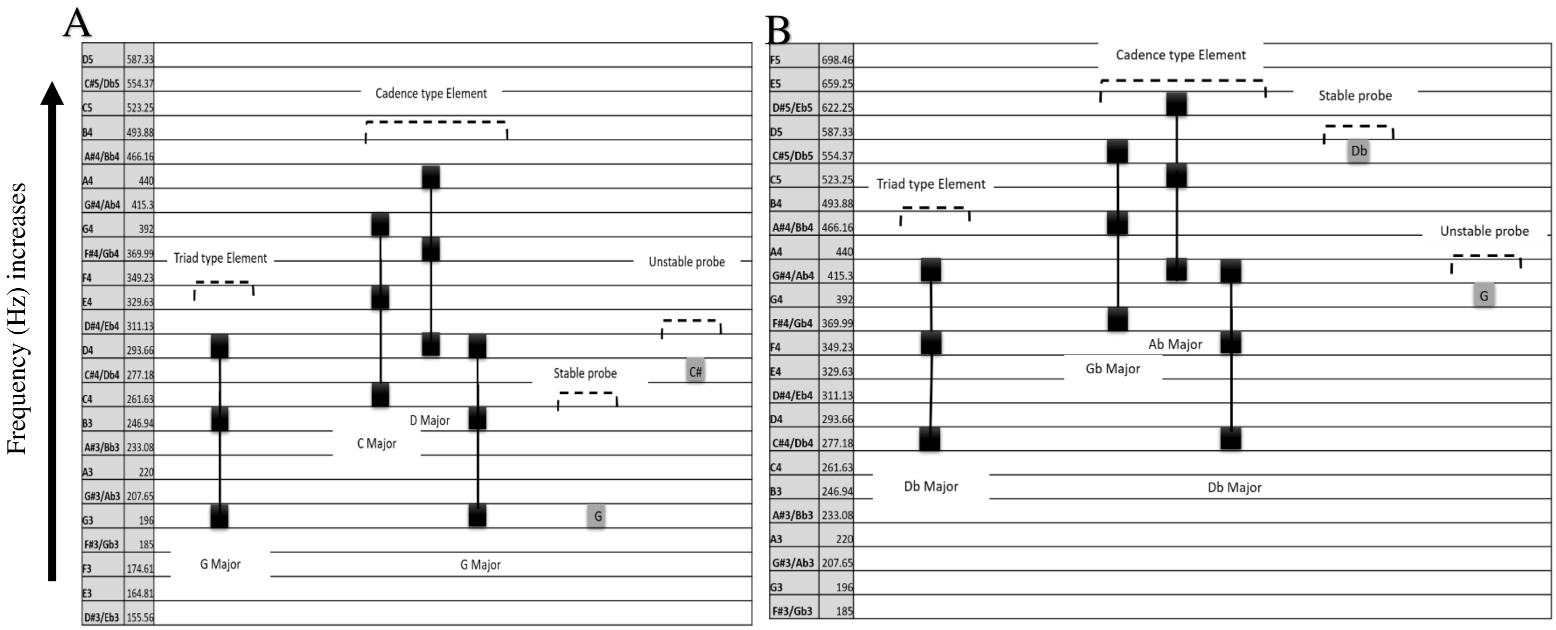
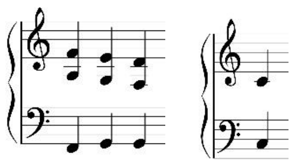


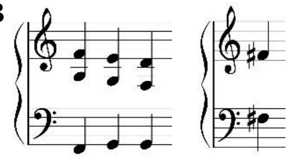
Figure S4. Schematic presentation of the auditory stimuli used in experiments 1, 4 & 4A (probe-tone method), in G Major (A) and in D flat Major (B). Each black box is a note, and the lines connecting the notes create a chord. The example shows a triad element and a cadence context element (a sequence of three chords), followed by an example of stable or unstable probe. In A, the elements are in G major key, the stable probe is the note G (the first scale degree, the most stable tone). The unstable probe is C# (the augmented forth degree, the tritone, which does not belong to the key, and is thus unstable). In B, the elements are in D flat major key, the stable probe is the note D flat (the first scale degree, the most stable tone). The unstable probe is G (the augmented forth degree, the triton, which does not belong to the scale, the least stable). Note that for the sake of clarity, each note in this schematic representation is represented by a single frequency. Stimuli used in this experiment, however, were Shepard tones [16], in which each pitch-class is created by 5 sine-tones separated by octaves.

Exp 2, 3 and 5

A



B



Audio S2. Examples of stable (A link) and unstable (B link) auditory stimuli used in Experiment 2, 3 & 5. Stimuli consist of a sequence of three chords followed by either a stable (A) or an unstable (B) tone.

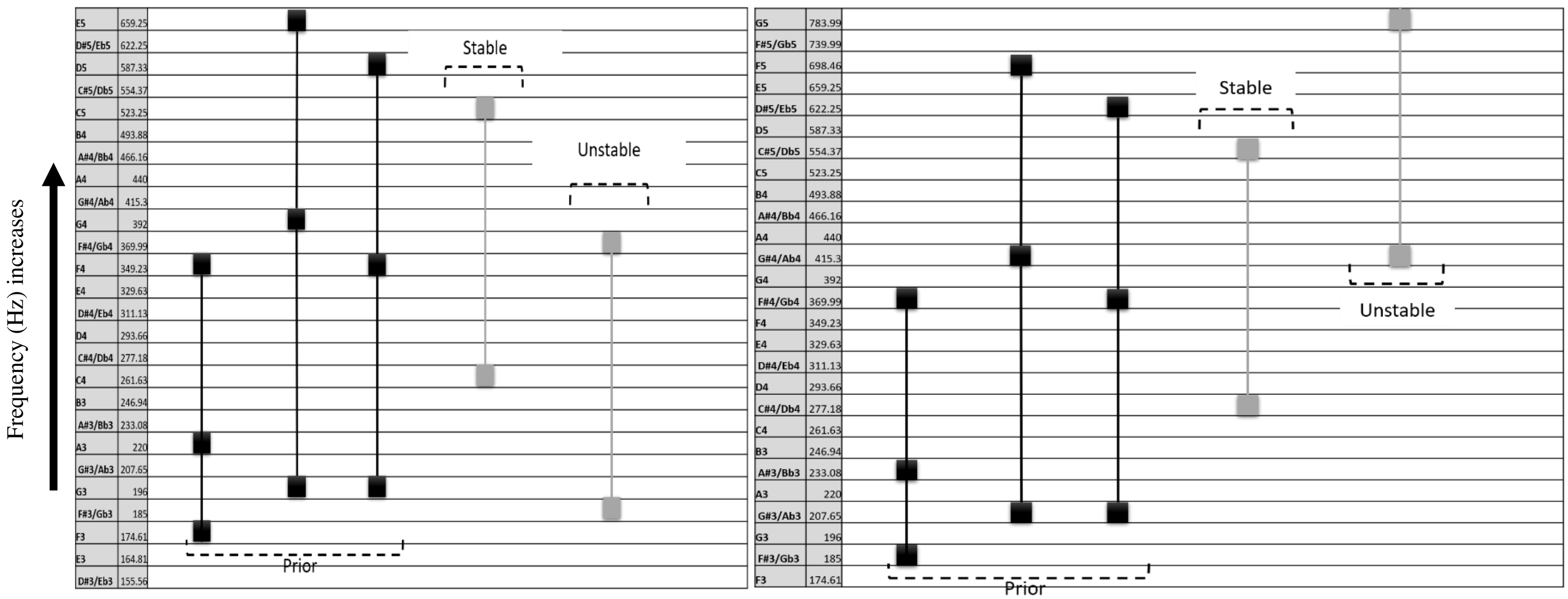


Figure S5. Schematic presentation of the auditory stimuli used in Experiment 2, 3 and 5 (IAT method), in C Major (A) and in C sharp Major (B). Each black box is a note, and the lines connecting the notes create a chord. The key in example A is C Major and consists of prior (a sequence of three chords), followed by a stable note (C, first degree of the scale, most stable) or an unstable note (F#, the raised 4th degree, least stable). The key in example B C sharp Major and consists of prior (a sequence of three chords), followed by a stable note (C sharp), first degree of the scale, most stable) and an unstable note (G, the raised 4th degree, least stable). Note that while for the sake of clarity each tone is represented here by its fundamental frequency (F0) only, stimuli consisted of harmonic tones (piano timbre).

References

1. Maimon N., Lamy D., Eitan Z., (2020) Crossmodal Correspondence Between Tonal Hierarchy and Visual Brightness: Associating Syntactic Structure and Perceptual Dimensions Across Modalities, *Multisensory Research*, **33**(8), 805-836.
2. E. Bigand, B. Tillmann, B. Poulin-Charronnat, A module for syntactic processing in music? *Trends Cogn. Sci.* **10**, 195–196 (2006).
3. M. Boltz, Perceiving the end: Effects of tonal relationships on melodic completion. *J. Exp. Psychol. Hum. Percept. Perform.* **15**, 749–761 (1989).
4. D. Huron, *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press, Cambridge, MA, USA (2006).
5. F. Lerdahl, C. L. Krumhansl, Modeling tonal tension, *Music Percept.* **24**, 329–366 (2007).
6. C. L. Krumhansl, Tonal hierarchies and rare intervals in music cognition, *Music Percept.* **7**, 309–324 (1990).
7. F. J. Fétis, *Traité Complet de la Théorie et de la Pratique de l'Harmonie Contenant la Doctrine de la Science et de l'Art*. (Maurice Schlesinger, Paris, France, 1844/2008).
8. A. Schoenberg, Harmonielehre (Leipzig:Verlagseigentum der Universal-Edition, 1911); trans. as *Theory of Harmony*, by Roy E. Carter (Berkeley: University of California Press, 1978).
9. J.P. Rameau. Traité de l'Harmonie Reduite à ses Principes Naturels: Divisé en Quatr Livres. (Jean-Baptiste-Christophe Ballard, Paris, France, 1722).
10. C. L. Krumhansl, The cognition of tonality–as we know it today. *Journal of New Music Research*, **33**, 253-268 (2004).
11. D. Shanahan, "Musical Structure: Tonality, Melody, Harmonicity, and Counterpoint". In *The Routledge Companion to Music Cognition*. M. Cobussen, V. Meelberg, B. Truax, Eds. (Routledge, NY, NY, USA, 2016), pp. 141-151.
12. J. C. Carlsen, Some factors which influence melodic expectancy, *Psychomusicology* **1**, 12–29 (1981).
13. L. L. Cuddy, C. A. Lunney, Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity, *Atten. Percept. Psychophys*. **57**, 451–462 (1995).
14. F. Marmel, B. Tillmann, Tonal priming beyond tonics. *Music Perception: An Interdisciplinary Journal*, *26*(3), 211-221 (2009).
15. R. Granot, E. Donchin, Do re mi fa sol la ti - constraints, congruity, and musical training: an event-related brain potentials study of musical expectancies, *Music Percept*. **19**, 487–528 (2009).
16. R. N. Shepard, Circularity in judgments of relative pitch. *The Journal of the Acoustical Society of America*, **36**, 2346-2353 (1964).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf