Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Pitch and time salience in metrical grouping

Jon Prince

School of Psychology, Murdoch University, Australia
j.prince@murdoch.edu.au

ABSTRACT

I report two experiments on the contribution of pitch and temporal cues to metrical grouping. Recent work on this question has revealed a dominance of pitch. Extending this work, a dimensional salience hypothesis predicts that the presence of tonality would influence the relative importance of pitch and time. Experiment 1 establishes baseline values of accents in pitch (pitch leaps) and time (duration accent) that result in equally strong percepts of metrical grouping. Pitch and temporal accents are recombined in Experiment 2 to see which dimension contributes more strongly to metrical grouping (and how). Both experiments test values in tonal and atonal contexts. Both dimensions had strong influences on perceived metric grouping, but pitch was clearly the more dominant. Furthermore, the relative strength of the two dimensions varied based on the tonality of the sequences. Pitch contributed more strongly in the tonal contexts than the atonal, whereas Time was stronger in the atonal contexts than the tonal. These findings are inconsistent with an interpretation that stimulus structure enhances the ability to extract, encode, and use information about an object. Instead, they imply that structure in one dimension can highlight that dimension at the expense of another (i.e., induce dimensional salience).

I. INTRODUCTION

How pitch and time combine in music perception has received renewed interest in the past few years (Barnes & Johnston, 2010; Boh, Herholz, Lappe, & Pantev, 2011; Boltz, 2011; Ellis & Jones, 2009; Firmino, Bueno, & Bigand, 2009; Henry & McAuley, 2011; Henry & McAuley, 2009; Johnston & Jones, 2006; Jones, Johnston, & Puente, 2006; Lebrun-Guillaud & Tillmann, 2007; Prince, 2011; Prince, Schmuckler, & Thompson, 2009; Prince, Thompson, & Schmuckler, 2009; Tillmann & Lebrun-Guillaud, 2006). Although earlier research largely characterised these dimensions as either independent or interactive (for a review, see Krumhansl, 2000), more recent work has sought to reconcile the conflicting findings and test systematically what factors influence the observed pattern of pitch-time combination (e.g., Boltz, 2011; Prince, Schmuckler, et al., 2009; Tillmann & Lebrun-Guillaud, 2006).

In my research I have explored the concept of dimensional salience (Prince, Thompson, et al., 2009), which corresponds to the prioritisation of a dimension in the mental representation of a stimulus, independent of perceptual difficulty (e.g., discriminability, see Garner, 1974). Reports of preferential emphasis on one dimension in perceptual processing exist in both auditory (Hébert & Peretz, 1997; Tong, Francis, & Gandour, 2008; Warrier & Zatorre, 2002) and visual domains (Atkinson, Tipples, Burt, & Young, 2005; Ellison & Massaro, 1997; Haxby, Hoffman, & Gobbini, 2000; Melara & Algom, 2003), despite matching the dimensions in terms of perceptual difficulty. Interpreted in the framework of dimensional salience, such reports show that a more salient dimension is more likely to interfere (i.e., demonstrate asymmetric interactions) with the processing of a less salient dimension. Depending on which dimension is more salient, the pattern of independence or interaction will likely vary accordingly. Specifically within typical Western music, pitch tends to be more salient than time (Prince, 2011; Prince, Schmuckler, et al., 2009; Prince, Thompson, et al., 2009), and possibly even in less typical music (cf. Ellis & Jones, 2009).

One of the most likely factors to realise dimensional salience is variation in the stimulus, however learned schemas that influence the perceptual processing of music may underlie this phenomenon. For Western listeners, the presence of a tonal context eliminates the effect of violations of temporal expectancies on pitch processing (Prince, Schmuckler, et al., 2009). That is, time was less salient (or pitch more salient) when sequences were tonal, but not when they were atonal. Without enculturation in a Western tonality there is no plausible reason for this result, thus it depended on the listener having learned the schema of tonality. Such effects may result from a gradual adjustment to the weighting of these dimensions based on the statistical properties of the stimulus.

In Western music, there is greater complexity in the dimension of pitch than time. For example, typical tonal melodies use 7 diatonic pitch classes (Järvinen, 1995; Knopoff & Hutchinson, 1983), use scale and contour as organisational principles (Dowling, 1978), and exhibit complex hierarchical structure in the form of tonality and harmony (Krumhansl, 1990). In contrast, 2-3 unique durations are the norm in this musical style (Fraisse, 1982), with a similarly elaborated metric organisation (Palmer & Krumhansl, 1990). Thus in terms of these statistical properties, pitch is clearly the more complicated dimension; in information-theoretic terms, this higher level of detail means it has more informative value. To optimise efficient processing of environmental stimuli, perceptual systems learn through experience to prioritise sources with more informative value (Goldstone, 1998), in vision (Bhatt & Quinn, 2011), speech (Werker & Tees, 2005), and music (Hannon, Soley, & Ullal, 2012). Accordingly, dimensional salience of pitch may reflect a priority of processing by being more useful for forming a mental representation of the stimulus.

Another potential source of dimensional salience is task design. Inherently temporal tasks such as tapping to a beat may highlight time over pitch (Pfordresher, 2003; Snyder & Krumhansl, 2001), whereas pitch-based tasks such as judging the goodness of a note or melody may favour pitch (Prince, 2011; Prince, Thompson, et al., 2009). Interestingly, Ellis and Jones (2009) found an overall dominance of pitch when measuring metrical grouping. These authors instructed listeners to rate on a 7-point scale the extent to which atonal sequences suggested metrical groupings of two notes (duple meter) or three (triple meter); the sequences used accents in pitch (pitch leaps) and time (duration differences) to suggest either duple or triple meter. Importantly, they equalised the strength of the dimensions by testing them separately prior to recombining them to create congruent or incongruent groupings (using values equalised in accent strength across dimensions). In these recombined sequences, pitch leap accents were by far the strongest factor, accounting for three times the variance of the temporal accents. Interpreted in terms of dimensional salience, pitch was clearly the more salient dimension even in this atonal time-based task, at least to the extent that metrical grouping can be considered an inherently temporal task (cf. Lerdahl & Jackendoff, 1983).

If the presence of Western tonality makes pitch more salient in music perception, then the strength of pitch accents versus temporal accents on metrical grouping may vary accordingly. Specifically, pitch accents should contribute more to metrical grouping in tonal sequences than they do in atonal sequences. Given that pitch was the more important dimension in the atonal sequences of Ellis and Jones (2009), pitch should always have a stronger contribution than time, but this advantage should be especially prevalent in tonal sequences. The goal of the current study was to test this possibility.

II. EXPERIMENT 1

In order to get an accurate assessment of dimensional salience in metric grouping, the dimensions (in this case, pitch and time) must be equated in difficulty in separate baseline conditions. Only then can any remaining difference between the dominance of dimensions be attributed to dimensional salience. Experiment 1 was conducted to establish equal strength of pitch and temporal accents in metric grouping (using pitch leaps and duration differences, after Ellis & Jones, 2009).

A. Participants

There were 12 participants in Experiment 1, with an average age of 29 years (*SD* = 11.5), not selected for their amount of formal musical training (*M* years of training = 1.3, *SD* = 2.5). All reported listening primarily to Western music throughout life and none reported a hearing deficit.

B. Stimuli

All stimuli were isochronous sequences (IOI = 500ms) of 24 notes, lasting in total 16 seconds. Sequences began with an additional three chords, none of which had the same duration as any of the subsequent notes; participants' instructions were to ignore these chords. In the tonal trials these chords were I-IV-V triads, and the subsequent sequence remained in the key of the first chord. According to the Krumhansl-Schmuckler keyfinding algorithm (Krumhansl & Schmuckler, 1986), the average correlation of tonal trials with the intended key was .87 (*SD* = .03). For the atonal trials, the chords did not establish any major or minor key, nor did the subsequent sequence (K-S correlation *M* = .46, *SD* = .06). There were two types of sequence, presenting varying levels of either a pitch leap accent or temporal duration accent.

1) Pitch accent stimuli

The sequences used to test the accent strength of pitch leaps were isochronous (IOI = 500ms, ISI = 250ms) ascending sequences, which had a pitch leap between either every two (duple meter) or three notes (triple meter) notes. Within a group (duple or triple), each subsequent note ascended to the next scale degree (e.g., C-D-E); between groups, the pitch leap accents consisted of ascending 2, 3, or 4 scale degrees (low, medium, and high accent strength, respectively). To remain in a single key, the intervals of the tonal trials could not always remain constant in their exact size. That is, sometimes the next scale degree was 1 semitone away, other times 2. Similarly, 2 scale degrees could be 3 or 4 semitones apart, 3 scale degrees could be 5 or 6 semitones, and 4 scale degrees could be 6 or 7 semitones. Thus the absolute pitch distance of a set scale degree change was not constant, potentially interfering with the nature of the accent, given that it was based on the size of the pitch leap. To ensure that this necessary feature of the stimulus design did not affect the tonal trials more or less than the atonal trials, the atonal sequences had the exact same number of each interval size. Table 1 shows a duple and triple example of a tonal and atonal sequence using a pitch leap accent of 2 scale degrees. For the duple examples, both tonal and atonal trials have 2 intervals of a semitone, 4 of two semitones, 2 of three semitones, and 3 of four semitones. In the triple examples, this distribution is slightly different from duple (2, 6, 2, and 1 instances of 1, 2, 3, and 4 semitones, respectively), but are still identical across tonal and atonal trials. In other words, the tonal and atonal trials differed by the sequential order of intervals, not the absolute size of intervals. Sequences started on either C3, C#3, D3, or D#3, ascended for 12 notes, and repeated.

2) Temporal accent stimuli

To ensure that both dimensions are equally powerful at inducing a metrical grouping before testing metrical grouping in sequences with both pitch and temporal accents, Ellis and Jones (2009) used a range of duration accents to induce temporal grouping, between 80 and 140ms. However, because these values are all below the temporal integration threshold of 200ms (Yost, 2000), they have a potential confound in that the listeners may have heard the duration accents as intensity accents (or a combination of both). In order to isolate the effect of duration accents, no notes were shorter than 200ms in the current studies.

Temporal accent stimuli were isochronous (equal IOI) but of unequal duration (unequal ISI). These accents created duration patterns of long-short (duple), or long-short-short (triple). The duration of (short) notes within a duple/triple group was 200ms. The first (long) note of the group, however, could be either 250, 333, or 450ms (low, medium, and high accent strength, respectively), corresponding to a Weber Fraction (WF) difference of either .25, .67, or 1.25. The WFs in Ellis and Jones (2009) were .33, .67, 1, or 1.25; thus the WFs in the current studies closely approximate those used previously, but at longer base duration (200ms instead of 80ms). These sequences were monotonic, but were still either tonal or atonal, on the basis of the chord sequence at the beginning of each trial. Sequences started on either C3, C#3, D3, or D#3, and remained on that pitch.

There were 2 dimensions (pitch, time), 2 types of metric grouping (duple, triple), 2 levels of tonality, 3 levels of accent strength, all with 4 different starting pitches, giving 96 unique conditions. Tonal and atonal trials were blocked and counterbalanced across participants. Trials were pseudorandomised such that the same grouping (duple or triple) would not be heard more than 5 times in a row. All participants heard all conditions twice, yielding 192 trials.

Table 1. Four example sequences (all starting on C) with group boundaries indicated with vertical lines (top – duple, bottom – triple). Each grouping type (duple, triple) and each accent size has a tonal (shaded) and atonal (unshaded) sequence. Both the notes (midi note naming) and sequential intervals are listed for each sequence. The frequency of interval sizes remains constant across tonality but varies across grouping type. In these examples the pitch leap accent is 2 scale degrees.

Grouping	Tonality		Notes and Intervals (semitones) in Sequence												Interval Frequencies: Within		Between	
															1	2	3	4
Duple	Tonal	Notes	C3	D3	F3	G3	B3	C4	E4	F4	A4	B4	D4	E4				
		Interval		2	3	2	4	1	4	1	4	2	3	2	2	4	2	3
	Atonal	Notes	C3	C#3	F3	G3	A#3	B3	D#4	F4	A4	B4	D4	E4				
		Interval		1	4	2	3	1	4	2	4	2	3	2	2	4	2	3
Triple	Tonal	Notes	C3	D3	E3	G3	A3	B3	D4	E4	F4	A4	B4	C4				
		Interval		2	2	3	2	2	3	2	1	4	2	1	2	6	2	1
	Atonal	Notes	C3	C#3	D#3	F#3	G#3	A#3	D4	E4	F#4	A4	B4	C4				
		Interval		1	2	3	2	2	4	2	2	3	2	1	2	6	2	1

C. Apparatus

Sequences were created as MIDI files in Finale Songwriter 2010, and then modified using custom MATLAB scripts to create all the necessary combinations, and then write to .wav files using a harmonically rich timbre. MATLAB scripts were also used to create the experimental interface using the PsychToolbox (Brainard, 1997), running on Windows XP. Participants wore Sennheiser HD280pro headphones to listen to each sequence.

D. Procedure

Trials were self-started by pressing the space bar, and participants indicated on a 5-point Likert scale to what extent the trials sounded like groups of two or three (1 = strong groups of two, 2 = slightly groups of two, 3 = unsure, 4 = slightly groups of three, 5 = strongly groups of three). Thus lower ratings correspond to perceptions of duple meter, higher ratings to triple meter. Participants performed 4 practice trials using sequences with congruent pitch and temporal accents, larger than those used in the full experiment for ease of demonstration. The experimenter was present for the practice trials to answer any questions, after which point the full experiment started with the participant alone in the quiet room. The entire procedure took approximately one hour.

III. RESULTS

Ratings were averaged across starting pitch (4) and repetition (2), giving 24 unique values for each participant. A 2 (dimension: pitch or time) x 2 (grouping: duple or triple) x 2 (tonality: tonal or atonal) x 3 (accent strength: low, medium, or high) repeated measures ANOVA using participants ratings of metric grouping revealed main effects of grouping, $F(1, 11) = 256.9$, $p < .001$, and accent strength, $F(2,22) = 4.3$, $p = .03$. These effects indicate that, reassuringly, participants rated duple sequences differently from triple sequences, but also that there was a slight difference among accent strength ($M = 2.89$, 2.99, 3.02) despite being averaged across duple and triple. None of the pairwise comparisons among accent strength levels were significant (again, averaged across grouping), so this effect is likely spurious.

There was also an expected interaction between grouping and accent strength, $F(2, 22) = 18.0$, $p < .001$. This effect indicates that participants rated the sequences differently across accent strengths for duple and triple groupings. That is, stronger accents lead to lower ratings (more groups of two) for duple groupings, but stronger accents lead to higher ratings (more groups of three) for triple groupings. Figure 1 depicts this pattern. No other main effects or interactions were significant.

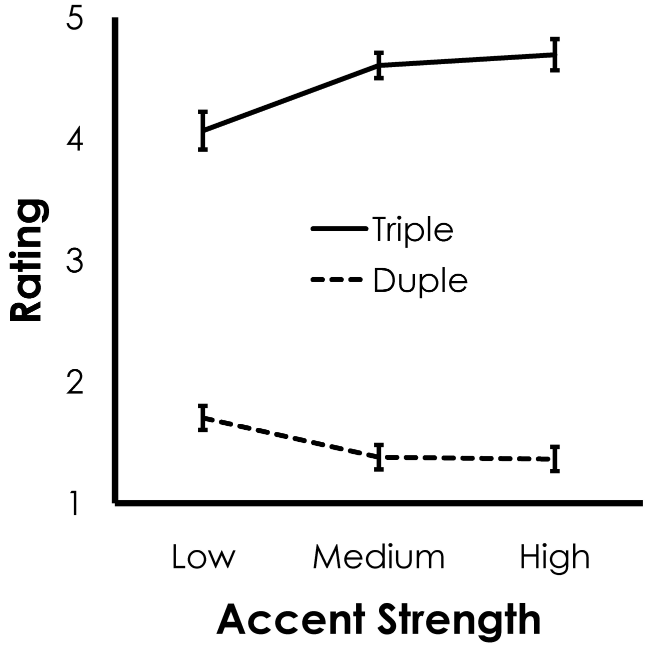

Figure 1. Metric grouping for pitch and temporal accents (averaged) in Experiment 1. Higher ratings mean triple percepts, lower ratings mean duple percepts.

IV. DISCUSSION

Tested separately, the selected accent strengths in both pitch and time successfully induced a perceived metric grouping matching the stimulus manipulation. Furthermore, the chosen values of accent strength were similarly effective for both pitch and time – that is, the ratings of strong pitch groupings did not differ from those of strong time groupings (and the same for the weak groupings). Additionally, the effects of the accent strength manipulation did not differ based on the tonality of the sequences, such that both pitch leap accents and temporal duration accents are equally powerful at inducing metric groupings in both tonal and atonal contexts. This result is especially convenient for the next experiment in which both pitch and temporal accents are present in individual sequences, because it means that simply varying the tonality does not alter the strength of the duple/triple manipulation. Rather, any change in effect size of pitch or time would signify that the relative emphasis on pitch or time changes based on the tonality of the sequence. Put differently, dimensional salience does not predict that temporal accents are ineffective in tonal contexts – when there are no competing pitch accents (in fact, there were no pitch accents at all). Instead, it predicts that when both pitch and temporal accents vary in the same sequence, the relative emphasis on one dimension will change if there is dimensional salience.

V. EXPERIMENT 2

The previous experiment tested the effectiveness of particular levels of pitch and temporal cues to metrical grouping, when presented in isolation, in both tonal and atonal contexts. Fortunately, it established that the chosen levels result in comparable metric grouping strengths across dimension. Experiment 2 aims to test if varying the tonality of the sequences will influence the weighting of pitch and temporal cues to metric grouping. That is, does a tonal context induce dimensional salience of pitch over time in measurements of metric grouping? There are three possibilities here – (1) the tonality of the context may have no effect whatsoever on the use of pitch and temporal cues; (2) additional stimulus structure (in this case, tonality) might aid the encoding, and thus effectiveness, of temporal cues (see Boltz, 1998; Jones & Boltz, 1989); (3) tonality may highlight the dimension of pitch, inducing dimensional salience of pitch which then results in weighting pitch cues more strongly than temporal cues. The first possibility is the obvious null hypothesis – pitch cues and temporal cues might be equally effective regardless of tonal context. The second eventuality would predict that temporal accents would be more powerful in tonal contexts, such that the absolute difference between low and high levels of temporal accent strength is larger for tonal contexts than atonal contexts. The third possibility makes the opposite prediction to the first – temporal cues should be weaker in tonal contexts than in atonal contexts, and vice versa for pitch cues, consistent with a dimensional salience hypothesis. Recombining the different levels of pitch and temporal accent strength allows testing of these possibilities, thus gaining insight into how pitch and temporal cues are combined in metric grouping.

A. Participants

A new set of 12 participants were in Experiment 2, with an average age of 28 (*SD* = 12.2) and 2.8 years of formal musical training (*SD* = 4.0). All reported listening primarily to Western music throughout life and none reported a hearing deficit.

B. Stimuli

Experiment 2 stimuli were constructed by recombining different levels of pitch and temporal accent strength tested in Experiment 1. Because there was no interaction between dimension, grouping, and accent strength in Experiment 1, pitch and temporal cues were equally strong at inducing a perceived metric grouping at each accent strength level. Nonetheless, the values with the closest nominal ratings across dimension were chosen for inclusion in the recombined versions. Specifically, pitch leap accents between groups were 5 scale degrees (accent strength level 3 from Experiment 1), and the temporal accents were 333ms between groups (accent strength level 2 from Experiment 1). Within groups, pitch intervals were 1 scale degree, and note durations were 200ms. A neutral metric grouping level was added to the existing duple and triple settings; the neutral grouping used a change every 6 notes, which is consistent with both a duple and triple metric grouping. Sequences started on the pitches C, C#, or D (omitting D#). Combining 3 levels of pitch accent groupings (duple, neutral, triple) with 3 levels of temporal accent groupings (duple, neutral, triple), 2 levels of tonality (tonal, atonal), and 3 starting pitches gives 54 unique stimuli. Participants heard all stimuli three times, resulting in 162 total trials.

C. Apparatus

The apparatus were the same as Experiment 1.

D. Procedure

The procedure was the same as Experiment 1; however with 30 fewer trials the experiment took a slightly shorter amount of time to complete.

VI. RESULTS

Ratings were averaged across starting pitch, giving 18 unique data points for each participant. A 3 (pitch grouping) x 3 (time grouping) x 2 (tonality) repeated measures ANOVA with participants' grouping ratings as the dependent variable revealed a main effect of pitch, $F(2, 22) = 102.7$, $p < .001$ and time, $F(2, 22) = 24.1$, $p < .001$, but not tonality, $F(1, 11) < 1$, *ns*. These effects indicate that both pitch and time groupings influenced participants' ratings, in the expected direction. Tonality interacted with both pitch and time, $F(2, 22) = 3.4$, $p = .05$, and $F(2, 22) = 4.4$, $p = .03$. These effects show that the effects of pitch and time differed for tonal and atonal contexts, and nearly so for the effect of time. Additionally, there was an interaction between pitch and time, $F(4, 44) = 14.0$, $p < .001$, because the effect of either dimension was larger (mean difference) when the other dimension suggested a triple grouping. In other words, triple groupings in either dimension were more susceptible to incongruity than duple groupings Figure 2 depicts this interaction. Lastly, the 3-way interaction between pitch, time, and tonality was not significant, $F(4, 44) = 1.5$, $p = .22$.

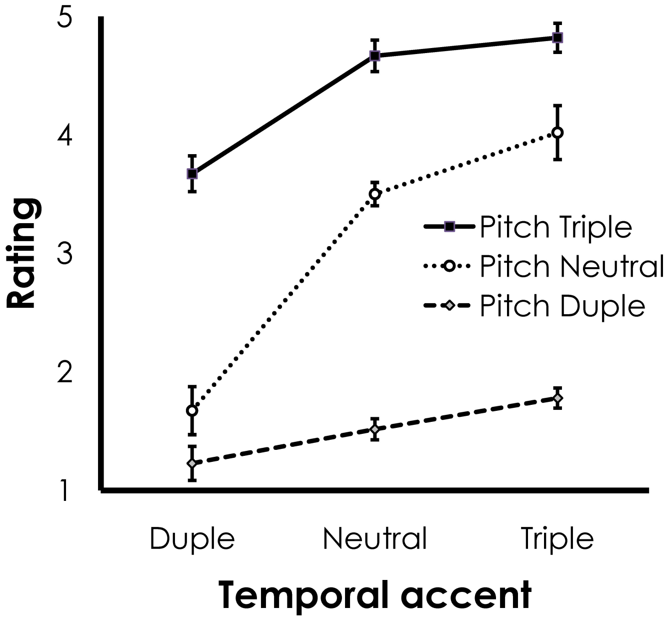

Figure 2: Ratings of metric grouping in Experiment 2, averaged across tonality to show the interaction between pitch and time. The interaction derives from duple groupings influencing the rating of neutral groupings (in the other dimension) more than triple groupings.

To explore the 2-way interactions between tonality, ratings were analysed separately for tonal and atonal trials with two 3 (pitch) x 3 (time) ANOVAs. Pitch, time, and their interaction had significant effects in both the tonal trials, $F(2, 22) = 71.9$, $p < .001$, $\eta^2 = .63$[1], $F(2, 22) = 18.2$, $p < .001$, $\eta^2 = .12$, $F(4, 44) = 8.8$, $p < .001$, $\eta^2 = .04$, and the atonal trials, $F(2, 22) = 114.4$, $p < .001$, $\eta^2 = .59$, $F(2, 22) = 29.5$, $p < .001$, $\eta^2 = .17$, $F(4, 44) = 12.3$, $p < .001$, $\eta^2 = .06$, respectively. The first result of note is that pitch effect sizes were always larger than time effect sizes. The second noteworthy finding comes from inspecting the change in effect sizes as a function of tonality. Pitch $\eta^2 = .63$ when the sequences were tonal, and time $\eta^2 = .12$; whereas for atonal sequences, pitch $\eta^2 = .59$, and time $\eta^2 = .17$. Put differently, the effect of pitch was 5 times the size of the time effect in tonal trials, and only 3.5 times the size of the time effect in atonal trials. Figure 3 shows how the influence of pitch and time on ratings varied as a function of tonality. It depicts difference scores representing the change in effectiveness of a dimension as a function of tonality for all participants (as well as the median). The difference scores were calculated from subtracting the average rating for the triple grouping trials from that of the duple grouping trials, for both the tonal and atonal contexts (separately). This difference in the atonal contexts was subtracted from the difference in the tonal contexts, resulting in a value representing how adding tonality to the sequences changed the weighting of the dimensions. Because pitch was more influential in tonal contexts, this difference score is positive; whereas time was less influential in these contexts, yielding a negative difference score. Not all participants showed this pattern (3 showed the opposite), but the median and mean are in the described direction. The final interesting finding was that the effect size of the interaction between pitch and time was slightly smaller for the tonal trials, $\eta^2 = .04$ than for atonal trials, $\eta^2 = .06$.

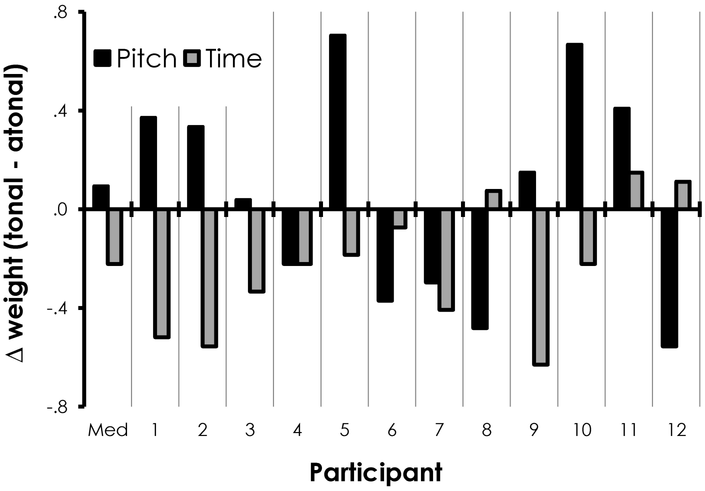

Figure 3: Change in the effectiveness of each dimension when comparing tonal and atonal trials for each participant (median on far left). Positive scores mean the dimension was weighted more in tonal trials, negative scores mean the dimension was less effective in tonal trials.

VII. CONCLUSION

These two experiments were designed to test if pitch and time contribute to the perception of metric grouping in tonal and atonal contexts. Experiment 1 was a test of the effectiveness of different levels of pitch and temporal cues (specifically, pitch leaps and duration accents) in establishing perceived metric groupings, for both dimensions separately. The goal of Experiment 2 was to test how pitch and temporal cues combined in tonal and atonal contexts; pitch cues accounted for much more variance than the temporal cues, even after having been equated in Experiment 1. That is, pitch was more salient than time – despite matching the effects of the two dimensions in separate baseline conditions, pitch was more influential when they were recombined. Lastly, not only were pitch cues more salient overall, they were more so when the sequences were tonal.

The data of Experiment 2 are inconsistent with the first two of the three possible scenarios explained earlier. First, tonality affected the use of pitch and temporal cues, rejecting the null hypothesis of no connection between tonality and dimensional salience. Second, adding pitch structure (i.e., tonality) did not improve the effectiveness of temporal cues, thus there was no evidence of tonality aiding the encoding and representation of temporal information. Instead, the data support the third possibility: tonality made the dimension of pitch more salient, thus magnifying the effect of pitch leap cues in metric grouping.

The finding that pitch was more influential than time on metric grouping is not new; Ellis and Jones (2009) reported an effect size of pitch 3 times that of time, importantly, using only an atonal context. For atonal contexts in Experiment 2, pitch

[1] Please note that all η^2 values are true eta-squared (not partial eta-squared).

accounted for 3.5 times as much variance as time, replicating well their results. Furthermore, the amount of variance accounted for by each dimension is also similar to the values reported by these authors. Specifically, the atonal contexts of the current experiment had a pitch η^2 of .59, and a time η^2 of .12; Ellis and Jones found pitch η^2 of .51 and time η^2 of .17. However, their statistically significant interaction between pitch and time was miniscule ($\eta^2 < .01$); in this experiment it was larger, both for the atonal ($\eta^2 = .06$) and tonal ($\eta^2 = .04$) contexts.

Indeed, across levels of tonality, both the effect sizes of the two dimensions and their interaction varied. When the context was tonal, pitch accents based on pitch leaps contributed more strongly than in atonal contexts. However temporal accents (duration cues) cues showed the opposite pattern – they affected metric grouping more strongly in atonal contexts than tonal contexts. Comparing the effect sizes across dimensions thus revealed an increased salience of pitch in the tonal trials (accounting for 5 times more variance) than in the atonal trials (3.5 times the variance). Thus, although pitch was always the more salient source of metric grouping cues regardless of the tonality of the contexts, it was especially so when the sequences were tonal.

Why would the interaction between pitch and time be stronger in the atonal context? One explanation is that pitch and time are more likely to interact when they are more comparably matched in terms of dimensional salience. In fact, this explanation is not new – in ratings of melodic goodness, Prince (2011) also found a variation in the strength of a pitch-time interaction, and proposed a possible explanation based on the relative size of main effects. Specifically, the size of the interaction was predicted well by the relative size of the main effects. When the difference between the main effect sizes was smaller (i.e., lower absolute value of pitch η^2 - time η^2), the interaction effect size was larger. In the present data, the effect sizes were closer in the atonal context (pitch η^2 - time η^2 = .42) than in the tonal context (pitch η^2 - time η^2 = .51), and the interaction was stronger in the atonal context (pitch * time η^2 = .06) than the tonal context (pitch * time η^2 = .04). Thus this pattern of effect sizes replicates that observed in Prince (2011).

Previous research into the use of temporal cues shows that increasing structure improves the encoding of the stimulus (Boltz, 1998; Jones & Boltz, 1989). However the present data show the opposite pattern – a tonal framework did not aid the use of temporal information, indeed it actually decreased its contribution. Boltz (1998) used typical Western melodies, and showed that coherent accent structure (congruent groupings implied by rhythm, rate, duration, and pitch interval sequence), and found that accuracy of reproduction and recognition of pitch changes was higher for structurally coherent sequences. Compared to the present experiments, her stimuli were more naturalistic – neither Experiment 1 nor 2 used sequences that would be described as typical melodies, not even when tonal. It is possible that by more closely resembling prototypical melodies, the schema involved in processing such melodies is better able to utilise structural coherence for encoding. Other work (Johnston & Jones, 2006; Jones, et al., 2006) demonstrates that listeners form expectancies based on both pitch and temporal structure, but only tested within an atonal context. However none of this work equated the strength of accents and structure from both dimensions, making difficult their interpretation in a dimensional salience context. The only research on pitch-time combination to meet this criterion is Ellis and Jones (2009) and Prince, Thompson, et al. (2009), but neither tested both tonal and atonal contexts, only one or the other. The present work clarifies these findings and extends the concept of dimensional salience to metrical grouping, arguably a strongly time-based task. In this sense dimensional salience seems relatively unaffected by task variations (for a similar result, see Prince, Schmuckler, et al., 2009), and more driven by the stimulus structure. In turn, this structure is dependent on listeners having internalised the schema of tonality. Activation of this schema leads to dimensional salience; possibly other schema in other modalities and domains may show similar effects on how stimulus dimensions combine.

REFERENCES

Atkinson, A. P., Tipples, J., Burt, D. M., & Young, A. W. (2005). Asymmetric interference between sex and emotion in face perception. *Perception and Psychophysics, 67*(7), 1199-1213.

Barnes, R., & Johnston, H. (2010). The role of timing deviations and target position uncertainty on temporal attending in a serial auditory pitch discrimination task. *Quarterly Journal of Experimental Psychology, 63*(2), 341-355. doi: 10.1080/17470210902925312

Bhatt, R. S., & Quinn, P. C. (2011). How does learning impact development in infancy? The case of perceptual organization. *Infancy, 16*(1), 2-38. doi: 10.1111/j.1532-7078.2010.00048.x

Boh, B., Herholz, S. C., Lappe, C., & Pantev, C. (2011). Processing of complex auditory patterns in musicians and nonmusicians. *PLoS ONE, 6*(7), 1-10. doi: e2145810.1371/journal.pone.0021458

Boltz, M. G. (1998). The processing of temporal and nontemporal information in the remembering of event durations and musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 24*(4), 1087-1104.

Boltz, M. G. (2011). Illusory tempo changes due to musical characteristics. *Music Perception, 28*(4), 367-386.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision, 10*, 433-436.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological review, 85*(4), 341-354.

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 264-280.

Ellison, J. W., & Massaro, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology-Human Perception and Performance, 23*(1), 213-226. doi: 10.1037//0096-1523.23.1.213

Firmino, É. A., Bueno, J. L. O., & Bigand, E. (2009). Travelling through pitch space speeds up musical time. *Music Perception, 26*(3), 205-209.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (1st ed., pp. 149-180). New York: Academic Press.

Garner, W. R. (1974). *The processing of information and structure* (Vol. 203). Oxford, England: Lawrence Erlbaum.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612. doi: 10.1146/annurev.psych.49.1.585

Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental*

Psychology: Human Perception and Performance. doi: 10.1037/a0027225

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223-233.

Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory & cognition, 25*(4), 518-533.

Henry, M. J., & McAuley, J. (2011). Velocity perception for sounds moving in frequency space. *Attention, Perception, & Psychophysics, 73*(1), 172-188. doi: 10.3758/s13414-010-0009-2

Henry, M. J., & McAuley, J. D. (2009). Evaluation of an imputed pitch velocity model of the auditory kappa effect. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 551-564.

Järvinen, T. (1995). Tonal hierarchies in jazz improvisation. *Music Perception, 12*(4), 415-437.

Johnston, H. M., & Jones, M. R. (2006). Higher order pattern structure influences auditory representational momentum. *Journal of Experimental Psychology: Human Perception and Performance, 32*(1), 2-17. doi: 10.1037/0096-1523.32.1.2

Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological review, 96*(3), 459-491.

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology, 53*(1), 59-96.

Knopoff, L., & Hutchinson, W. (1983). Entropy as a measure of style: The influence of sample length. *Journal of music theory, 27*(1), 75-97.

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

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological bulletin, 126*(1), 159-179.

Krumhansl, C. L., & Schmuckler, M. A. (1986, July). *Key-finding in music: An algorithm based on pattern matching to tonal hierarchies.* Paper presented at the 19th Annual Meeting of the Society of Mathematical Psychology, Cambridge, MA.

Lebrun-Guillaud, G., & Tillmann, B. (2007). Influence of a tone's tonal function on temporal change detection. *Perception and Psychophysics, 69*(8), 1450-1459.

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

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110*(3), 422-471.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception & Performance, 16*(4), 728-741.

Prince, J. B. (2011). The integration of stimulus dimensions in the perception of music. *Quarterly Journal of Experimental Psychology, 64*(11), 2125-2152. doi: 10.1080/17470218.2011.573080

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009). The effect of task and pitch structure on pitch-time interactions in music. *Memory and Cognition, 37*(3), 368-381. doi: 10.3758/MC.37.3.368

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance 35*(5), 1598-1617. doi: 10.1037/a0016456

Tillmann, B., & Lebrun-Guillaud, G. (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research, 70*(5), 345-358.

Tong, Y. X., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes, 23*(5), 689-708. doi: 10.1080/01690960701728261

Warrier, C., & Zatorre, R. (2002). Influence of tonal context and timbral variation on perception of pitch. *Attention, Perception, & Psychophysics, 64*(2), 198-207. doi: 10.3758/bf03195786

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental psychobiology, 46*(3), 233-251.

Yost, W. A. (2000). *Fundamentals of hearing: An introduction* (4th ed.). San Diego, CA, US: Academic Press.