Proceedings of the 9th International Conference of Students of Systematic Musicology (SysMus16), Jyväskylä, Finland, 8th - 10th June 2016. Birgitta Burger, Joshua Bamford, & Emily Carlson (Eds.).

Isomorphism of Pitch and Time

Olivia Xin Wen[1], Carol Lynne Krumhansl[1]

[1]Psychology, Cornell University, USA
xw383@cornell.edu

Abstract

An ongoing debate regarding the perception of pitch and time is whether information on the two dimensions is processed independently or interactively. To study this, we tested whether listeners prefer sequences in which tonally stable tones coincide with rhythmically stable tones. Our study builds on a noted isomorphism between pitch intervals in the diatonic scale and tone durations in the standard rhythm originating in Ghana. This isomorphism is shown in a) the maximally even structure of 2212221 and b) the cyclic nature with seven possible starting points. To better understand pitch-time relationship, we conducted two experiments. In Experiment 1, we created seven scales based on the diatonic pattern and seven rhythms based on the standard pattern by shifting the starting pitch interval or tone duration. To measure the perceived tonal stability of tones in the scales, in Experiment 1a each scale was followed by a probe tone and listeners judged how well the tone fit into the scale. To measure the perceived rhythmic stability of tones in the rhythms, in Experiment 1b each position of the sequences was accented dynamically and listeners judged how well the accent fit into the rhythm. These ratings were then used in analyzing the results of Experiment 2 that used all 49 pairs combining the 7 scales and 7 rhythms in Experiment 1. Participants rated a) how well the rhythm fits the scale for each pair and b) familiarity and well-formedness of each scale and rhythm. Results show that probe ratings from Experiment 1 predict judgments in Experiment 2. Specifically, scale/rhythm pairs received higher ratings when tonal and rhythmic hierarchies correlated more strongly with each other. In addition, we found a familiarity bias toward the major scale. After accounting for this bias, results remain significant, suggesting that information from the two individual dimensions interact perceptually.

Keywords: pitch-time relationship, diatonic scale, standard pattern, tone duration, probe tone

Introduction

Pressing (1983) noted the isomorphism shown in Figure 1a between the pitch intervals in the diatonic scale in Western music and the durations in what is called the standard pattern. Whereas the diatonic scale is the most prominent scale in Western tonal music, the standard pattern is the most used African rhythmic pattern, originating in Ewe dance music in Ghana (Agawu, 2006). It can be found running in the background of Afro-Latin music, such as Salsa. The Standard Pattern is often played in a repeated fashion, meaning that, once the rhythm ends, it loops back to the beginning. This repetition can also be found in the domain of scales; it takes the form that the scale intervals repeat when the octave is reached. In fact, this cyclic feature is only one of the parallels between the diatonic scale and the standard pattern. As Figure 1b shows, this shared structure between scale and rhythm can be explained as a maximally even set (Clough & Douthett, 1991), where the seven white circles are maximally evenly spread among the twelve evenly distributed positions around a circle. On the circle, the distance between every two adjacent positions on the circle can represent a semitone in pitch or an eighth note in duration. Starting from the white circle at the 12 o'clock position and counting clockwise, the distance between every two adjacent white circles can be expressed as 2212221, which can be called the diatonic pattern. This structure describes the pitch intervals between every two consecutive tones in the diatonic scale, and it also describes the temporal intervals between every two consecutive temporal positions in the standard pattern.

The diatonic pattern has two important features. One is that it is asymmetrical (Browne, 1981; Rahn, 1996). Each of the tones bears a unique constellation of relations to every other tone. So even if listeners only hear two or three of them, they can still tell where in this pattern they are, and where the most stable tone or the most stressed tone is. As briefly touched on earlier, the other feature is that the diatonic pattern is cyclic (Iyer, Bilmes, Wright, & Wessel, 1997; Temperley, 2000). This allows a scale or a rhythm to start from any of these seven points and then cycle back to the starting point, which generates seven unique patterns. This theoretically prominent isomorphism between the diatonic scale and the standard pattern suggests a new approach to studying how pitch and time combine perceptually.

In the ongoing debate regarding how information about pitch and time combine, a number of studies have provided evidence for two opposing positions (as summarized in Krumhansl, 2000; see also Prince & Schmuckler, 2014). One is that pitch and time are two separable dimensions, where one dimension does not interact with the perception of the other. In support of this, the perception of melodic similarity has been found to be an additive function of the similarity of the melodic patterns and the similarity of the rhythmic patterns (Monahan & Carterette, 1985). A similar result was found in a study finding that judgments of phrase endings were an additive function of tonal and metrical hierarchies (Palmer & Krumhansl, 1987a&b). Another source of support for the independence position comes from the neuropsychological literature in which patients may loose sensitivity to melodic information while retaining the ability to distinguish between rhythms (e.g., Peretz & Kolinsky, 1993).

The other position is that pitch and time interact perceptually. This means that change in one dimension affects the perception of the other. A corpus study showed that tonally stable pitch classes tend to occur at temporally stable positions (Prince & Schmuckler, 2014). Other studies found a pitch bias where tonal stability affects judgments of temporal positions (Prince, Thompson, & Schmuckler, 2009) and meter perception (Ellis & Jones, 2009). In the opposite direction, studies have found better memory for tones occurring at rhythmically expected points in time (Jones, Moynihan, MacKenzie, & Puente, 2002). Overall, the precise nature of the pitch-time relationship is not yet well understood.

Inspired by the isomorphism between the diatonic scale and the standard pattern, the current experiments took an alternative approach to the question of how pitch and time are processed. The experiments asked whether judgments of how well a rhythm fit a scale could be accounted for by how much the tonally stable tones and the rhythmically stable tones coincided. That is, were the fit judgments higher when the two hierarchies of stability correlated with one another?

Experiment 1

The purpose of this experiment was to measure tonal and rhythmic stability in the seven scales and seven rhythms that are formed by shifting the starting pitch interval or tone duration in Figure 1b. In order to measure the perceived tonal stability of the tones in the scales, Experiment 1a was a probe tone experiment in which each of the probe tones following the scales was judged as to how well it fit with the scale context. In order to measure the perceived rhythmic stability of each of the tones in the rhythms, in Experiment 1b each position of the sequences was accented dynamically and listeners judged how well the accent fit with the rhythm. These judgments were then used in the analysis of the results of Experiment 2 that used all 49 possible combinations of the 7 scales and 7 rhythms in Experiment 1.

Method

Participants

Forty-five Cornell University students participated in each experiment for course credit or a $5 cash reward. Thirty-one participated in both. In both studies, all but 1 were musically trained. Participants in experiment 1a had an average of 14 years of musical training; 3 had absolute pitch. Participants in experiment 1b had an average of 12.8 years of musical training.

Stimulus materials

In both experiments, all sequences consisted of 8 tones, forming 7 intervals. They were created using GarageBand and were sounded in piano timbre. In Experiment 1a, the seven scales were constructed by shifting the starting interval on the diatonic pattern in Figure 1b. For example, Scale1 had the intervals (in semitones) 2212221, Scale2 had the intervals 2122212, and so on. Each scale was constructed in both ascending and descending forms, beginning and ending on C. The range of the ascending sequence was C2 to C3, and the range of the descending sequence was C4 to C5 (details see Krumhansl & Shepard, 1979). The seven scales were followed by a probe tone that was one of the seven tones in the scale, which was played in the range of C3 to C4 between the ranges of the two scale contexts. A baseline trial was also composed for each scale with isochronous rhythm. For Experiment 1b, we constructed seven rhythms by shifting the starting duration on the standard pattern in Figure 1b. For example, Rhythm1 had the durations (in eighth notes) 2212221, Rhythm2 had the durations 2122212, and so on. They were played monotonically on C3. On successive trials, each of the seven temporal positions was dynamically accented; this is called the probe accent. A baseline trial was also composed for each rhythm with monotonic pitch. All rhythms were played twice with a short pause in between.

Procedure

Participants were asked to listen to one stimulus at a time and then make their rating. In Experiment 1a, they rated how well the probe tone fits into the scale by moving a slider on a continuous scale from extremely bad fit to extremely good fit. In Experiment 1b, they rated how well the probe accent fits into the rhythm by moving the same slider. In both studies, they first completed four practice trials. The trials were blocked by scale or rhythm. Each block began with the relevant scale or rhythm played in a neutral form, that is, without a probe tone or accent. The neutral form for Experiment 1a was an ascending scale followed by the same scale in descending order, with a short pause in between; the neutral trial for Experiment 1b was a rhythm played twice with a short pause in between. After the neutral trial, they listened to and rated the probes for that particular scale or rhythm, and then moved on to a different scale or rhythm which were presented in a randomized order. Once they rated all probe trials for the seven scales or rhythms, they listened to and rated each of the 7 baseline trials in a randomized order on two scales: how familiar the scale or rhythm is, and how well-formed the scale or rhythm seems (in other words, whether the scale or rhythm forms a good pattern). Both items were also rated by moving a slider on a continuous scale, from extremely unfamiliar or ill-formed to extremely familiar or well-formed. At the end of the study, they filled out a demographics questionnaire. Each study lasted approximately 30 minutes.

Experiment 2

Method

Participants

Fifty Cornell University students participated in the experiment for course credit. All but 4 were musically trained. Participants had an average of 11.3 years of musical training; 3 had absolute pitch.

Stimulus materials

Forty-nine scale/rhythm pairs were constructed by combining the seven scales and the seven rhythms used in Experiment 1. As Table 1 shows, out of all 49 stimuli, the seven on the diagonal in this table are matched pairs, because both the scale and the rhythm start from the same point in the diatonic pattern in Figure 1b. This means that the scale and the rhythm share the same structure. The rest are mismatched pairs, because the scale and the rhythm do not share the same structure. In addition, the same fourteen baseline trials from Experiment 1 were used again as baseline trials.

Procedure

Participants were asked to listen to one stimulus at a time. After each listening, they rated how well the rhythm fits the scale by moving a slider on a continuous scale from extremely bad fit to extremely good fit. First, they completed four practice trials, and then rated the 49 pairs in a randomized order. After filling out a demographics questionnaire, they listened to and rated the 49 pairs for a second time, in a different randomized order. After filling out another demographics questionnaire, they listened to and rated familiarity and well-formedness for each of the 14 baseline trials, presented in a randomized order. The study lasted approximately 30 minutes.

Results

Data were processed in the following way. All continuous rating scales were coded from -100 to 100, with -100 being extremely bad fit, unfamiliar, or ill-formed, and 100 being extremely good fit, familiar, or well-formed. For Experiment 1a, probe tone ratings from ascending and descending trials were averaged because they correlated highly with each other. Next, individual ratings were averaged across participants because no large effect of musical training background was found in either Experiment 1a or 1b. This way, one judgment rating was obtained for each probe tone in each scale, and one judgment rating was obtained for each probe accent in each rhythm. Probe tone ratings was then correlated with probe accent ratings, which gives a predicted goodness of fit measure for how participants would judge the combined pitch and time pattern in Experiment 2. Table 2 shows this goodness of fit measure from Experiment 1.

Similarly, in Experiment 2, the two judgment ratings for each scale/rhythm pair were averaged because they correlated highly with each other. Individual ratings were also averaged across participants because no large effect of musical training background was found. Table 3 shows the ratings of how well the rhythm fit the scale for each of the 49 pairs from Experiment 2. To determine how information about pitch and time combine perceptually, we correlated the goodness of fit measure from Experiment 1 with the judgment ratings from Experiment 2.

If pitch and time are separable dimensions, then the exact scale/rhythm combination should not matter for the cross-dimension judgment. In other words, the correlation between probe tone ratings and probe accent ratings should not predict the cross-dimension judgment ratings. Thus, the expected correlation between the goodness of fit measure and the judgment ratings would be zero if the two dimensions are processed separately. On the other hand, if the correlation is not zero and is significant, then it means that pitch and time are not separable dimensions and that they interact in perception.

Results show a positive and significant correlation between the goodness of fit measure and judgment ratings, $r(49) = .65$, $p < .0001$. This suggests that pitch and time interact in the perception of music. They interact in such a way that listeners prefer the higher-rated tones to be played on higher-rated temporal positions.

However, listeners in Experiment 2 reported being much more familiar with the major diatonic scale than the others. Consequently, we computed the residuals of listener's judgments after taking out the effect of familiarity. A correlation analysis was conducted to assess the judgments against the goodness of fit measure. Results remain positive and significant, $r(49) = .51$, $p < .001$. This suggests that after taking the familiarity bias into account, listeners still preferred the higher-rated tones to be played on higher-rated temporal positions.

In addition, the judgment ratings were examined against the surface-level structural match between pitch interval and tone duration. As Table 1 shows, all matched pairs were coded as 1 and the rest as 0. Correlation between judgment and surface-level match was not significant, $r(49) = .05$, $p = .75$. The surface-level match was also coded in two other ways. One way was to count the number of times the pitch interval and the time interval matched. The other was how many positions matched before the mismatch. Neither of the codings of surface-level match correlated significantly with the judgments of how well the rhythm matched the scale; for the first coding, $r(49) = .03$, $p = .86$; for the second coding, $r(49) = .17$, $p = .23$. This suggests that the surface-level structural match does not predict judgments of cross-dimension fit. Instead, it is the match between the underlying tonal stability and rhythmic stability of the tones that predicts judgments of fit.

Discussion

The current experiments explored the relationship between musical pitch and time by focusing on the isomorphism between pitch interval and tone duration. Specifically, scales and rhythms in the diatonic pattern were used as stimuli. Findings suggest that the surface-level structural match did not predict judgments of the cross-dimension fit. Instead, the correlation between the two probe ratings, measuring tonal stability and rhythmic stability, predicted judgments of how well the rhythm fit the scale. The ratings were higher when the higher-rated tones were played on the higher-rated temporal positions in the probe experiments. This suggests that listeners' cross-dimension judgments were governed by their preference for the best-fitting tones in the diatonic scales to be played on the best-fitting temporal locations in the standard pattern. This finding shows that pitch and time are not two separable dimensions. Instead, they interact when joined together.

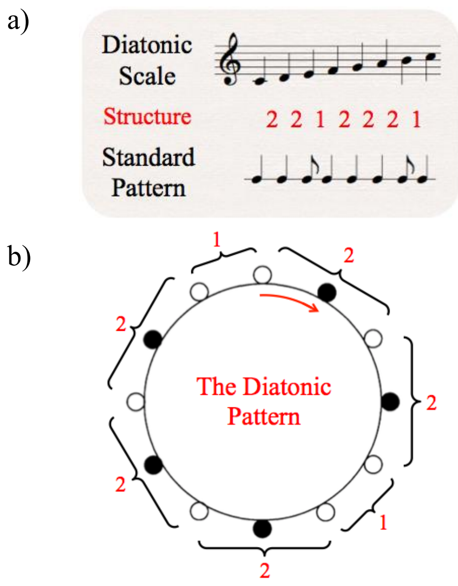

Figure 1. a). Isomorphic structure shared by the pitch intervals of the diatonic scale and the tone durations of the standard pattern. b). Illustration of the asymmetrical and cyclic Diatonic Pattern.

	Scale1	Scale2	Scale3	Scale4	Scale5	Scale6	Scale7
Rhythm1	1	0	0	0	0	0	0
Rhythm2	0	1	0	0	0	0	0
Rhythm3	0	0	1	0	0	0	0
Rhythm4	0	0	0	1	0	0	0
Rhythm5	0	0	0	0	1	0	0
Rhythm6	0	0	0	0	0	1	0
Rhythm7	0	0	0	0	0	0	1

Table 1. Design for Experiment 2 and surface-level coding

Note. Matched pair = 1, mismatched pair = 0

Table 2. Goodness of fit measure constructed by correlating probe tone ratings and probe accent ratings

	Scale1	Scale2	Scale3	Scale4	Scale5	Scale6	Scale7
Rhythm1	.90	.62	.52	.86	.71	.76	.12
Rhythm2	.77	.51	.49	.68	.64	.68	-.05
Rhythm3	.30	.07	.59	.20	.21	.30	.29
Rhythm4	.56	.33	.25	.37	.50	.47	-.17
Rhythm5	.86	.64	.47	.73	.76	.75	.01
Rhythm6	.33	-.16	.22	.45	.06	.08	.14
Rhythm7	.42	.03	.17	.23	.36	.19	-.08

Table 3. Judgments ratings of cross-dimension fit

	Scale1	Scale2	Scale3	Scale4	Scale5	Scale6	Scale7
Rhythm1	27.61	0.39	1.40	4.87	4.61	1.39	-15.95
Rhythm2	41.16	1.81	0.13	9.88	6.55	8.57	-5.83
Rhythm3	13.43	-13.44	-11.24	0.11	-7.96	-9.01	-24.63
Rhythm4	16.69	-0.34	-3.55	12.62	7.00	-5.50	-6.81
Rhythm5	39.27	-0.67	0.17	10.34	6.51	6.30	-9.11
Rhythm6	32.60	-12.54	-5.25	4.71	-9.57	-0.35	-12.99
Rhythm7	10.04	-20.08	-21.19	-7.29	-21.41	-10.13	-24.13

Note. The lowest possible value is -100, meaning extremely bad fit; the highest possible value is 100, meaning extremely good fit.

Acknowledgment

We would like to thank the SEMPRE Conference Award and the Cognitive Science Travel Grant at Cornell University in supporting Olivia Wen for her attendance of this conference. We would also like to thank Matthew Shortell for making stimuli and collecting data for Experiment 1.

References

Agawu, K. (2006). Structural analysis of cultural analysis? Competing perspectives on the “Standard Pattern” of West African Rhythm. *Journal of the American Musicological Society, 59,* 1–46.

Browne, R. (1981). Tonal implications of the diatonic set. *In Theory Only, 5,* 3–21.

Clough, J., & Douthett, J. (1991). Maximally Even Sets. *Journal of Music Theory, 35,* 93–173.

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

Iyer, V., Bilmes, J., Wright, M., & Wessel, D. (1997). A novel representation for rhythmic structure. *In Proceedings of the 23rd International Computer Music Conference,* 97–100.

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

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126,* 159–179.

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 579–594.

Monahan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception, 3,* 1–32.

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Jounal of Experimental Psychology: Human Perception and Performance, 13,* 116–126.

Palmer, C., & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics, 41,* 505–518.

Peretz, I. & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *The Quarterly Journal of Experimental Psychology, 46A,* 301–325.

Pressing, J. (1983). Cognitive isomorphisms between pitch and rhythm in world musics: West Africa, the Balkans and Western Tonality. *Studies in Music, 17,* 38–61.

Prince, J. B., & Schmuckler, M. A. (2014). The tonal-metric hierarchy: A corpus analysis. *Music Perception, 31,* 254–270.

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,* 1598–1617.

Rahn, J. (1987). Asymmetrical ostinatos in sub-Saharan music: Time, pitch, and cycles reconsidered. *In Theory Only, 9,* no.7: 23–36.

Temperley, D. (2000). Meter and grouping in African music: A review from music theory. *Ethnomusicology, 44,* 65–96.