

A STUDY OF EXPLORATORY ANALYSIS IN MELODIC SONIFICATION WITH STRUCTURAL AND DURATIONAL TIME SCALES

Takahiko Tsuchiya and Jason Freeman

Georgia Institute of Technology,
Center for Music Technology,
840 McMillan St. Atlanta GA 30332, USA
{takahiko, jason.freeman}@gatech.edu

ABSTRACT

Melodic sonification is one of the most common methods of sonification: data modulates the pitch of an audio synthesizer over time. This simple sonification, however, still raises questions about how we listen to a melody and perceive the motions and patterns characterized by the underlying data. We argue that analytical listening to such melodies may focus on different ranges of the melody at different times and discover the pitch (and data) relationships gradually over time and after repeated listening. To examine such behaviors in real-time listening to a melodic sonification, we conducted a user study employing interactive time and pitch resolution controls for the user. The study also examines the relationships of these changing time and pitch resolutions to perceived musicality. The results indicate a stronger general relationship between the time progression and the use of time-resolution control to analyze data characteristics, while the pitch resolution controls tend to have more correlation with subjective perceptions of musicality.

1. INTRODUCTION

Consider asking someone to sing a melody based on one-dimensional, numeric, and sequential data (e.g., an EEG sensor input). Although there may be various mapping possibilities with the human voice, many people may choose to modulate the tone (pitch) of their voice by looking at the approximate value changes of the data sequence. Now, consider listening to such a pitch-modulated melody and trying to analyze the characteristics of the underlying data, such as recurring shapes that might be short, long, or both. Although these attributes might depend largely on the sonified data itself, are there any common patterns in how we explore and identify unknown data attributes over time?

First, we need to ask how our mind processes a newly-encountered melody in general. To understand the structure of a melody in real time, our mind does not seem to simply follow the instantaneous pitches from moment to moment. Instead, depending on the analytical goals, we may compare the current pitch to the immediate past, several seconds ago, to the beginning (as much as working memory helps), or possibly even to the upcoming (i.e., “anticipated” [1]) pitches to search for the characteristics of the melody. In other words, in an auditory analysis of data within a melody, our listening “focal range” over time may change dynamically.

This study examines our general exploratory behaviors when listening to and analyzing the data attributes in local and global time spans in a melodic sonification. Considering the obvious difficulty of measuring variable focal levels of listening, we have instead conducted an experiment featuring a generative system with dynamic resolution controls for the listener¹. With the control sliders, the user can physically emulate the varying focal ranges in real time by abstracting (e.g., quantizing) the data in time and pitch dimensions.

Our primary goal was to observe and statistically test how the listener explores and understands the structure of an unfamiliar melody, and how their analysis develops over time as the melody unfolds and repeats. However, the experimental design was also informed by two additional interests: to examine the relationships of the emulated resolution parameters to a subjectively-chosen musical balance, as the notion of time structures has a strong tie to musical aesthetics; and to evaluate the advantages of an exploratory real-time sonification system that allows the alternation of the system behavior.

We expect that this study contributes to the development of sonification techniques in many ways. It may, for example, inform the parameter-tuning process of a simple melodic sonification for optimal analytical experiences without expending additional mapping channels in the audio synthesis. It may also help creating a more engaging and intuitive sonification by taking a “balanced” resolution mapping for musical and analytical interests according to our test results. Lastly, by employing active control of parameters for users rather than a typical passive listening environment, we hope to present unique opportunities as well as challenges in analyzing a highly interactive sonification.

2. RELATED WORK

The perception and treatment of time structures in a melody involve various philosophical and pragmatic discussions. Particularly in the field of contemporary music, where digital audio synthesis has enabled “sub-symbolic” manipulations of sounds, the study of time structures in sound has been popular among many composers and theorists. Here, we briefly address their work and concepts most relevant to our experiment in two broad categories: ones for musical (qualitative) and cognitive perspectives and ones for more quantitative perspectives related to (functional) sonification, visualization, statistics, and digital signal processing. These opposing categories, however, often intersect with each other.



This work is licensed under Creative Commons Attribution – Non Commercial 4.0 International License. The full terms of the License are available at <http://creativecommons.org/licenses/by-nc/4.0/>

¹ <http://quantizationtest.herokuapp.com>
(User-experiment website, Retrieved: 3/13/2018)

2.1. Multiplicity of Musical Time

The common consensus in the study of musical time is that sounds are expressed and perceived not simply in the absolute “clock” time but in multiple ways. Desain and Honing, for instance, observe that a typical musical expression consists of, at minimum, discrete time intervals (e.g., meters) and continuous and “expressive” timing manipulations (e.g., “rubato” and “accelerando”), which poses a challenge in automatic transcriptions [2]. While these competing timing factors are both in the note- or phrase-level time durations, various theorists consider musical time to be hierarchical, even for a simple monophonic melody. Roads describes nine levels of time scale that contribute to the formation of musical piece, sections, rhythms, pitch, timbre, etc. for both perception and generation [3]. Here, different time scales imply different resolutions or levels of detail. The sensation of tone (pitch), for instance, happens at the “sound object” (i.e., note) and “microsonic” levels when there are enough repetitions of a microsonic pattern, while timbral effects are noticed when there are “micro” and “sample” level irregularities. The melodic patterns are recognized in the “meso” time scale, one level above the sound objects. How would these structural time scales or resolutions play a role when analyzing the data within a melodic sonification, such as value-by-value fluctuations, local and global peaks, and gradual shifts of the central point? How does the listener identify these scales in a new melody?

In addition to the hierarchical time scales, our experience of musical time is multiple while listening to and analyzing a melody, as expressed as “polyphony of viewpoints” by Jonathan Kramer [4]. He observes that when we listen, learn, and compare different parts of the melody, our mind does not simply follow the absolute time with a linear progression. Instead, we learn the structure of the melody in terms of the duration of auditory patterns and their various proportions (cf. time scales). As we listen to the piece / melody, Kramer observes, we acquire new information about the proportions. As such, the focal length of time may be constantly adjusted with two cognitive processes present: one following the durations in passing (from a still sounding past moment until the current moment), and the other experiencing and comparing the remembered durations in retrospect. While we may continuously learn the time structure of the melody as it progresses, he also argues the importance of repeated listening for a thorough analysis [5].

The notions of nonlinearity and multiplicity in musical time, therefore, have informed our user-experiment designs, where we attempt to capture the experience of absolute time vs. the duration relative to the total performance time of the melody (including repetitions), take the progress of time into consideration as the process of learning the periodic structures, and encourage the repeated listening experience where the user may learn the optimal focal ranges according to the current position (phase) of the data.

2.2. Measurable Contents in Absolute Time

The measurement of the listener’s response to multiple time scales is, however, a complex task. Pragmatic sonifications tend to rely on the uniformity of absolute time with a well-defined and steady clock. Here, instead of the varying qualities in time scales, the experience of time may be altered by the content or the data within the absolute time scale. This may involve various strategies in mapping and transforming data to sound, including the use of statistical abstraction techniques.

In this regard, the audification technique maps the individual data points to successive samples, with the assumption that our ability of listening to and processing the audio-rate high-density data is as efficient as the short-time Fourier transform, with the time and frequency-domain resolutions high enough to perceive the timbral details [6]. Such assumption may, however, become questionable for the perceptibility of sample-level movements, especially with data with irregular motions, even though the global characteristics might be effectively captured as a spectral / timbral impression.

To counter the uncertainty of the playback speed, many melodic sonifications such as the Sonification Sandbox [7] utilize the auditory “tick marks,” or a static metronome, to indicate the current time position of the playback in the data. This approach, however, does not necessarily capture the potential hierarchical structure in a melodic sequence, or might even introduce a false sense of periodicity. The mapping of data to the tempo can also be variable and arbitrary, potentially affecting the perception of the melodic structures.

Our experiment takes an alternative approach where we present the melody with a fixed tempo, but instead offer statistical abstraction tools to alter the mapping of data contents to a constant (absolute) time line.

Data visualization, as well as sonification, frequently employ statistical aggregation or abstraction techniques that effectively translate continuous data to discrete symbols. For example, the histogram of continuous data groups values into discrete “bins” with uniformly-spaced quantization steps [8]. Scaler or vector quantization are used ubiquitously for digitization [9], structural analysis and compression [10]. While various researchers in sonification warn against quantizing signals to create “music-like” effects [11], nonlinear transformations, including quantization, are an important tool for analyzing and handling both continuous and symbolic data, especially ones with discrete and hierarchical structures [12].

3. USER EXPERIMENT

In this interactive listening experiment, we collected anonymized user data from 20 Georgia Tech Music-Technology students who, according to the survey, had a wide range of listening and musical-performance skills from novice to professional.

3.1. Goals and Hypotheses

In this experiment, we asked subjects to perform analytical tasks utilizing the dynamic time and pitch resolution controls for aiding and measuring their structural analysis process. While we hoped to find general correlations between their “final” parameters and the optimal perception of certain data attributes, such results may be also dependent on the data source and, ultimately, the user’s personal preference. Instead, our greater interest was to find patterns in their exploratory behaviors on a newly-encountered sonification, measured over several different durational time scales. In other words, our study questions if a progress of time correlates the way they like to hear the melody in certain time / pitch resolutions for a particular analytical goal.

We have hypothesized that, taking Kramer’s argument of the listener gaining new information about melodic proportions through the progress of time and repetitive listening, there may be converging behaviors or common directionalities for both time- and pitch-resolution adjustments

over time. We also questioned if the total duration spent for a task has a correlation to their final values.

As for the use of the time / pitch resolution controls that emulate the structural analysis of listening, Roads' hierarchical time-scale theory suggested to us that creating an appropriate pitch resolution may be more important to capture the phrase-level and longer time-scale structures of the melody, such as the global contour, than for local details. On the other hand, the time-resolution control may have stronger relationships especially to the perception of local-level data attributes, such as small peaks and rapid fluctuations.

3.2. The System for Experiment

The recent introduction of the Web Audio standard¹ has enabled us to develop real-time audio synthesis for all major web browsers. Utilizing web audio, we have previously developed an application programming interface (API) called data-to-music (DTM) [13] that facilitates rapid prototyping and embedding of sonification into any web applications [14]. This online user study utilizes the DTM API to create a real-time user-customizable sonification.

3.3. The Dataset

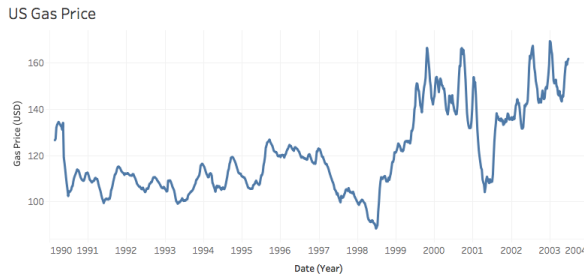


Figure 1: The dataset for the experiment: the US Gas Price time series.

For every test, we used the US Gas Price dataset², single-dimensional data sampled at a fixed time interval from 1990 to 2003 with 695 data points in total (Figure 1). We mapped the gas-price time series to a monophonic oscillator voice using the square waveform, with a logarithmic frequency modulation (FM) that scales the data domain of [88.5, 169.3] (USD) to a MIDI-note range of [45, 110]. The data was played back for the duration of 17.5 seconds, with about 39.7 data points per second at the default time / pitch resolution. (Though not strongly intended, 40Hz is used typically for speech analysis and modeling as the minimum threshold for the perception of tone [3].) The amplitude of the melody stays constant. In the experiment, the user does not see the visualization of the data, but only hears the sonified result.

3.4. User Control of Resolutions

The test provides a minimal user interface consisting of the Play / Stop buttons, a playback-position indicator, a time-resolution control slider, and a pitch-resolution control slider. The resolution controls modify the melody being played in real time by altering the mapping of the underlying data.

¹ <https://webaudio.github.io/web-audio-api> (Retrieved: 3/13/2018)

The pitch-resolution control applies a uniform but variable-step-size scaler quantization onto the log-scaled (i.e., MIDI-note) frequency. As the user lowers the pitch resolution, for example, the output MIDI range is rounded by the factor up to 9.0.

The time-resolution control down-samples and interpolates the data. It first segments the data into a smaller number of blocks (with the minimum of 30), calculates the mean value for each block, then interpolates the result into the original length of the data (695). We assumed a linearly-interpolating characteristic between data points (as in a line graph) rather than a step interpolation (as in a bar graph) and rendered the pitch-modulated sound accordingly. The use of the linear interpolation is perhaps the most debatable design decision we took. This is because, when mixed with the pitch quantization in some tasks, the effects of step-interpolated data points would become rather indistinguishable from the pitch quantization.

When the time and pitch abstractions are combined (e.g., in tasks 0, 3, 6, and 7, described in detail in the following section), the pitch quantization is applied after the time abstraction (interpolation) process. Both time and pitch resolution controls are recorded in the normalized range of [0, 1], and set to 1 (max resolution and unaltered) as the default value for each task.

3.5. User Analytical Tasks

Table 1: Task types, numbers and enabled controls.

Type / Ctrl	Time Res.	Pitch Res.	Both
Practice	---	---	0
Details-Local	1	2	3
Local-Global	4	5	6
Musical	---	---	7

The main purpose of the experiment was, again, to observe the subject's exploration over time as they perform an analytical task on a melodic sonification. However, defining an analytical listening task is itself a challenge. We cannot, for instance, ask to identify a specific pattern in the melody without giving away the information about the data and biasing their exploratory behaviors. Instead, the tasks ask the user to look for a "balancing" point or a perceptual boundary between two high-level categories of data attributes that are most likely in different time spans.

We categorize the target attributes into the time spans of "details," "local," and "global" levels. The "details" level includes perceived attributes such as rapid fluctuations or continuities of individual data points. The "local" time span may include small peaks and periodic patterns, as well as the current central tendencies. The "global" attributes may be such as the overall contour of the data, the total value range, and general directions of the data sequence. The analytical tasks are grouped into two main stages: Tasks 1 through 3 ask the user to find a boundary between the "details" and "local" characteristics, and tasks 4 through 6 ask to identify the balancing point between "local" and "global" levels. Each task features different combinations of the time / pitch resolution controls (see Table 1).

The task 0 was presented as the practice stage in order for the user to get familiar with the behaviors of the control UIs. The task 7 is an extra stage that asks to find the most musically-balanced resolutions without the consideration for data analysis.

² <https://vincentarelbundock.github.io/Rdatasets/datasets.html> (File: gasprice.csv; Retrieved: 3/13/2018)

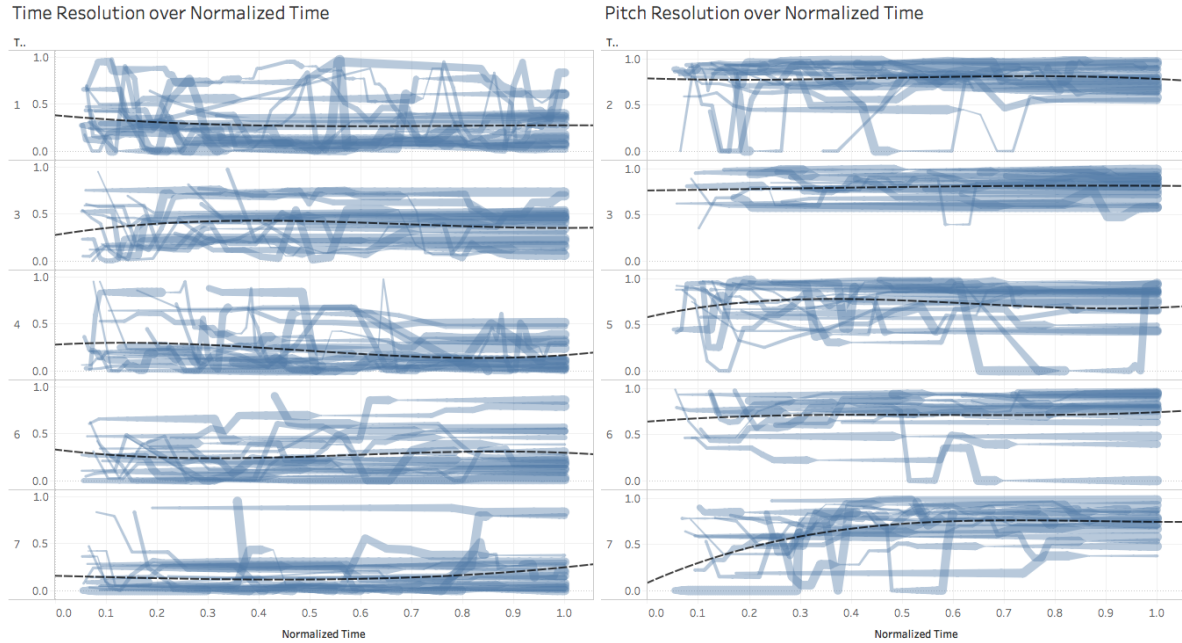


Figure 2: Time / pitch resolution control over the normalized time. The thickness of the line indicates the current phase of the data. The dotted lines are the 3rd order polynomial regression models. Note that some values are filtered out for unbiasing the regression models (see section 4.1).

3.6. Measurements and Abbreviations

The experiment records the user input of the time resolution (TR) and pitch resolution (PR) controls. These user inputs were polled every second with the current time stamp and playback position (phase) of the data. Not only the relationships between TR and PR, our analytical interests are their relationships to various scales of time, including the tasks {1, 2, ..., 7}, absolute time (AT), normalized time (NT), and data phase.

AT, measured in seconds, provides the information about the total duration the user spent until reaching the final value of the resolution parameters for each task. This may include repeated playback of the data. AT, however, cannot be used for directly comparing the user inputs over time.

NT scales the AT spent in each task to the range of [0, 1], where the TR or PR value at NT = 1 is the final value chosen for the task. With NT, we can compare the users' exploratory / adjusting behaviors over the duration of the entire listening.

The data phase, recorded in [0, 1] range, is the third time scale regarding the current playback position of the melody. It encodes the information about the data source itself, where we may potentially find exploratory trends corresponding to a particular region in the data.

In the following analyses, “TR or PR” is denoted as “TR/PR” as a shorthand. The final resolution values at NT = 1 are denoted as “fTR/fPR.”

4. ANALYSES

4.1. Preprocessing

Every user task starts with the default TR/PR set to 1 (i.e., no alternations to the data). This has resulted in artificially correlated tendencies near the starting point of each task. To reduce the biases, for TR/PR vs. NT, we excluded the data

points that were TR/PR > 0.999 (i.e., values most likely unchanged by the user) as well as NT < 0.1. Similarly, for the histograms and TR/PR vs. phase, we filtered out the data when TR/PR > 0.999 as the default-value bias. Since the phase cycle can repeat multiple times per task, we did not filter out their starting points where phase < 0.1.

4.2. Regression Models

The statistical analyses utilize linear or third-order polynomial (for the user data over the timeline that tend to fluctuate) regression models, computed with the ordinary-least-squares method in Tableau 10.5. In the model statistics, the R^2 describes the variance of the recorded data points from the prediction model, with higher values (~ 1.0) being a better fit to the model. The P-value shows the probability that the output is affected by random chance (i.e., the model coefficients are set zero), with a lower value (e.g., $p < 0.05$) indicating the statistical significance of the input variable. Many of the regressive models below use the time (e.g., NT or phase) as the input, and TR/PR as the predicted output.

4.3. Overall Results

In general, the regression models with any input (e.g., time) to TR/PR outputs show very low R^2 values, commonly below 0.1 (10%), suggesting that the trend lines cannot be used for a precise prediction of the “exploratory” analysis. We speculate the factors to be either the general random nature of user exploration, strong personal biases of perceived structures in the melody, or potentially the ineffectiveness of the task instructions (see section 4.4). On the other hand, some models have the p-value below 0.01, indicating the presence of common directionalities of two variables when compared to the case where the input coefficient is set to zero (i.e., a null hypothesis; see sections 4.5 and 4.6). The model directionalities are, however, generally moderate as seen in Figures 2 and 7.

4.4. Analyses with Aggregate Time

Before analyzing the time-evolving aspects of the subjects' data exploration, we examined the aggregate TR/PR for each task disregarding the time progress. The results generally suggest strong personal preferences of TR as the final values.

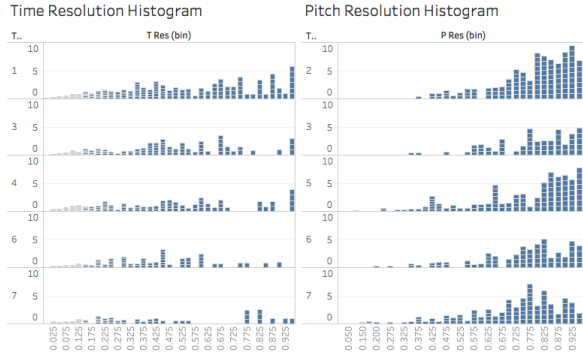


Figure 3: The histogram (20 bins) of the average TR/PR for each task. The frequencies were normalized by the total AT for each user.

We first looked for the regularities in TR/PR in each analytical task, which may indicate commonly-agreed structural resolutions when disregarding the time progress. Here, a normal or sharp distribution may indicate a value range for the optimal perception of structures. The results (Figure 3) show some trends in PR-only controls (tasks 2 and 5) and task 7 (“musical”). TR, however, show high spreads and little trends. Note the filtered values (see section 4.1) affecting the TR results in tasks such as 6 and 7 – indicating that multiple users preferred to mostly alter PR for those tasks.

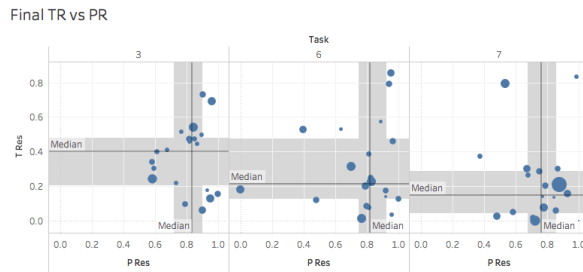


Figure 4: The final TR vs. PR values for tasks 3, 6, and 7. The size of dot indicates the amount of time spent. The gray areas contain the 2nd and 3rd quartiles.

Are there any proportional patterns with fTR and fPR? The tasks 3 (details-local), 6 (local-global), and 7 (musical) take the combined controls of TR and PR. The comparisons (Figure 4) show almost no correlations between fTR/fPR. It shows the shifting median for fTR but not fPR. The stationary nature of fPR can be explained with regard to subjective musical preferences (see section 4.6).

Table 2: The linear regression model statistics between fTR/fPR (outputs) and task transitions (inputs).

Ctrl	TR				PR			
	1→4	3→6	3→7	6→7	2→5	3→6	3→7	6→7
R ²	0.052	0.022	0.122	0.036	0.001	0.008	0.036	0.003
P-val	0.156	0.356	0.026	0.237	0.844	0.575	0.234	0.736

Next, we used the linear regression to model the general directionalities of the final values between different tasks. While the aggregate fTR looked to have a linear trend (see Figure 4), the individual fTR values between tasks had few common directionalities. Table 2 shows that only fTR for the tasks 3→7 were statistically significant, implying some time-scale relationships between “details-local” and “musical,” but more general evidence of strong personal preferences.

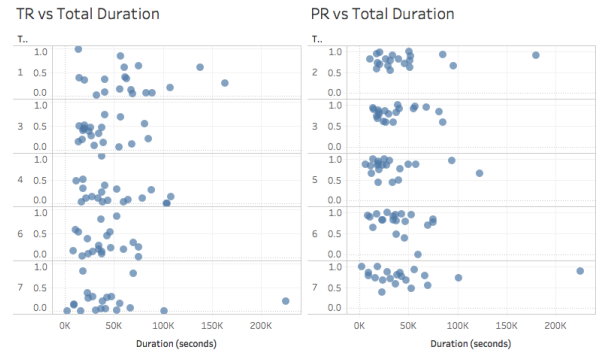


Figure 5: The final TR/PR values vs. the total time spent for each task.

Lastly for the aggregate time analyses, we examined the effect of the total duration spent vs. the final values. All results, however, showed little to no correlation between these variables (Figure 5). The number of repetitions also did not uniformly affect the final values. These results may signify the widely-different personal preferences among the subjects for the use of TR/PR for analysis.

4.5. Analyses with Time Progress

While there appears to be little commonly-agreed fTR/fPR that may explain the melodic or data structure, are there any regularities in how the subject explores the data before arriving the final values? Although the user explorations have certainly complicated the analyses of the result, we observed several characteristic outcomes. In order to directly compare the activities among users, we scaled the recorded data into a normalized time line (NT), and also mapped to the current phase of the data.

4.5.1. Explorations over Normalized Time

Table 3: The 3rd-order polynomial regression model statistics for TR/PR (outputs) vs. NT (input).

TR	1	3	4	6	7
R ²	0.010	0.018	0.079	0.013	0.038
P-value	0.002	0.006	< 0.0001	0.024	< 0.0001

PR	2	3	5	6	7
R ²	0.005	0.007	0.026	0.003	0.242
P-value	0.231	0.225	0.001	0.609	< 0.0001

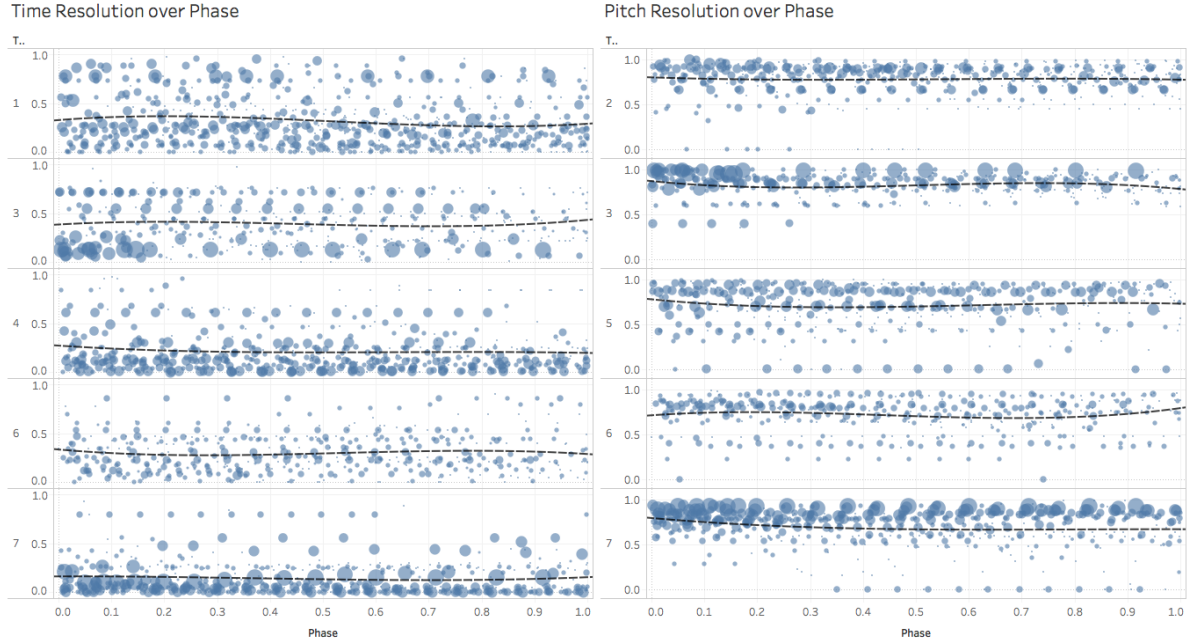


Figure 6: The TR/PR controls over the data phase. The circle size indicates the replay count.

In the normalized time line for each task, the regression analysis (Table 3; Figure 2) indicates that there are common directionalities in the user exploration affected by the time progression. Particularly, the subjects seem to utilize the TR control for exploring the data across all tasks. PR, on the other hand, only had correlations to time in the PR-only local-global structure analysis (5) and the musicality task (7). This suggests the limitation of PR for analyzing particular attributes, such as the sample- or local-level details.

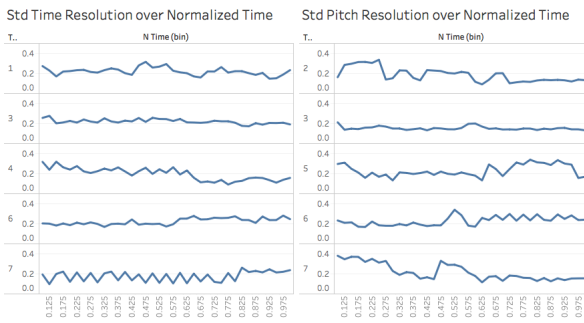


Figure 7: The standard deviation of TR/NR over NT.

Would the time progression directly affect the choice of a commonly-accepted resolution value? Figure 7 shows the amount of convergence (standard deviation) toward a specific resolution value over time. The results, however, did not indicate such converging behaviors except for PR in tasks 2 and 7 and TR in task 4. This may suggest the stronger influence of individually-controlled TR on the local-global analysis while PR has more influence on the details-local analysis as well as the musical-balancing task.

4.5.2. Explorations over Data Phase

Do the subjects react similarly to any “points of interest” in data? To examine this, we compared their exploratory paths with regard to the phase of data, where, within the total duration of NT or AT, each user may repeat the playback and adjust the previously-set resolutions.

Table 4: The 3rd-order polynomial regression model statistics for TR/PR (outputs) vs. phase (input).

TR	1	3	4	6	7
R^2	0.022	0.006	0.005	0.006	0.006
P-value	< 0.0001	0.460	0.360	0.401	0.353

PR	2	3	5	6	7
R^2	< 0.001	0.023	0.007	0.018	0.029
P-value	0.928	0.037	0.324	0.060	0.001

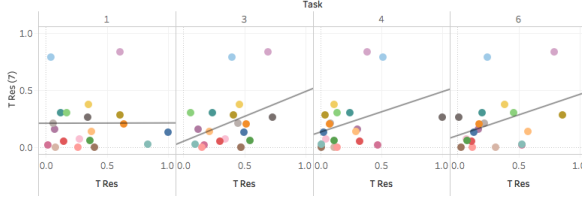
The results (Figure 6, Table 4) are, to our surprise, almost contrary from the case of NT, that TR shows much less common directionalities than PR with regard to the phase. We consider that these high variance and random nature may indicate several possibilities:

- As the combined controls (tasks 3, 6, and 7) indicate the significance of phase vs. PR over TR, the PR is more useful in location-sensitive analysis.
- TR may be used more than PR for random-directional explorations. TR may be first used as the base line for PR adjustments in NT time scale, and PR is used for micro adjustments. As such, the results may be reflecting the “replay” factor of phase where each iteration of TR may start and progress from a much different base line.
- The TR control is not suitable for an “immediate” analysis sensitive to the phase. Or, this particular dataset or melodic configuration (e.g., speed) does not guide the TR analysis very well.

4.6. Analytical Balance vs. Musical Balance

Lastly, we observed the relationships between the final resolution values chosen for each task and subjective preferences of musicality using those controls.

TR analysis vs TR musical



PR analysis vs PR musical

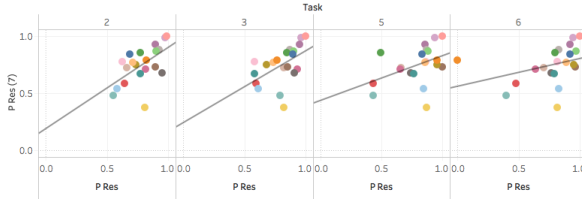


Figure 8: fTR/fPR (inputs) vs. fTR/fPR in task 7 (output). The users are color coded.

Table 5: The linear regression model statistics for fTR/fPR (inputs) vs. fTR/fPR in task 7 (outputs).

Ctrl	TR				PR			
Tasks	1	3	4	6	2	3	5	6
R ²	<0.1	0.126	0.128	0.142	0.342	0.265	0.178	0.130
P-val	0.989	0.123	0.121	0.101	0.006	0.020	0.063	0.117

Figure 8 and Table 5 compare the analytical fTR/fPR with the musical fTR/fPR in task 7. Again, the final values are “static” and does not capture the time-varying explorations. However, as the analytical task moves from details-local to local-global, the TR seems to become more correlated whereas the PR slightly decorrelates. This may also confirm that the individual choices of fPR for analytical tasks are generally related to their musical preferences, thus stay stationary across tasks (see section 4.4). The fTR, in contrast, does not generally correlate with musical preferences, especially for details-local time-scale analysis.

5. DISCUSSIONS

In this study, we observed several contrasting patterns in how the listener explores and matches the TR/PR controls to the structures of the melody (data). In the task level, where we disregard the time progression, we saw strong correlations between the musical preferences and all pitch configurations, whereas the TR varied for both the perception of data and musical balances. Over the time line of data exploration, on one hand, the TR had much larger effects in all tasks until arriving the final configurations. This defied our hypothesis that TR/PR would affect separately on local and global structures. On the other hand, in repeated listening analyses over phase, the PR had more significance in the exploration. These suggest the durational time scales for listening (i.e., separate tasks, normalized time, and phase) may utilize different combinations of the resolution controls, but not as simple as time to the low-level and pitch to the high-level relationships.

The results also showed that the final and overall motions of the resolution parameters were highly divergent, perhaps reflecting the personal preferences for listening and analysis.

In this regard, it is important to note that, while the use of the time / pitch abstraction techniques may bring forward different time-scale structures in the melody to our attention, using a lower resolution physically filters out the detailed motions from the currently-produced sound. Thus, it should be encouraged to actively change and scan through different resolutions than assuming the final optimal value early in the exploration.

6. CONCLUSIONS

In this study, we examined the use of dynamic time and pitch resolution controls for listening and identifying various attributes of a melodic sonification. We argued the considerations in multiple hierarchical structures in a melody, and how to possibly measure them. Our experiment also emphasized the exploratory natures of listening and understanding a newly-encountered melody. Unlike a passive listening-and-response test, our interactive user test provided a real-time sonification with user customization and recording. This presented challenges in analyzing their exploratory approaches, which were highly variable and complex. However, the unique relationships that were found between different time scales and musical preferences may bring new design opportunities for creating engaging and dynamic melodic sonifications.

7. REFERENCES

- [1] D. B. Huron, *Sweet Anticipation: Music and the Psychology of Expectation*. MIT press, 2006.
- [2] P. Desain and H. Honing, “The Quantization of Musical Time: A Connectionist Approach,” *Comput. Music J.*, vol. 13, no. 3, pp. 56–66, 1989.
- [3] C. Roads, *Microsound*. Cambridge: MIT Press, 2001.
- [4] J. D. Kramer, “Multiple and Non-Linear Time in Beethoven’s Opus 135,” *Perspect. New Music*, vol. 11, no. 2, pp. 122–145, 1973.
- [5] J. D. Kramer, *The Time of Music: New Meanings, New Temporalities, New Listening Strategies*. 1988.
- [6] R. L. Alexander, J. A. Gilbert, E. Landi, M. Simoni, T. H. Zurbuchen, and D. A. Roberts, “Audification as a Diagnostic Tool for Exploratory Heliospheric Data Analysis,” Jun. 2011.
- [7] B. N. Walker and J. T. Cothran, “Sonification Sandbox: A Graphical Toolkit for Auditory Graphs,” in *Proc. ICAD*, 2003, vol. 3.
- [8] F. J. Gravetter and L.-A. B. Forzano, *Research Methods for the Behavioral Sciences*. Nelson Education, 2015.
- [9] P. R. Cook, “Sound Synthesis for Auditory Display,” *Sonification Handb. Logos Publ. House Berl.* 2011.
- [10] T. Painter and A. Spanias, “Perceptual Coding of Digital Audio,” *Proc. IEEE*, vol. 88, no. 4, Apr. 2000.
- [11] T. Bovermann, J. Rohrerhuber, and A. de Campo, *Laboratory Methods for Experimental Sonification*. Berlin: Logos Verlag, 2011.
- [12] R. Rowe, “Sub-Symbolic Processes,” in *Machine Musicianship*, MIT press, 2001, pp. 93–144.
- [13] T. Tsuchiya, J. Freeman, and L. W. Lerner, “Data-to-Music API: Real-Time Data-Agnostic Sonification with Musical Structure Models,” in *Proc. ICAD*, 2015.
- [14] R. M. Winters, T. Tsuchiya, L. W. Lerner, and J. Freeman, *Multi-Modal Web-Based Dashboards for Geo-Located Real-Time Monitoring*. Georgia Institute of Technology, 2016.