How Predictable Do We Like Our Music? Eliciting Aesthetic Preferences With The Melody Triangle Mobile App

Henrik Ekeus[1], Samer A. Abdallah[2], Peter W. McOwan[1], Mark D. Plumbley[1]
[1]Centre for Digital Music, Queen Mary University of London
[2]Department of Computer Science, University College London
{hekeus,peter.mcowan,mark.plumbley}@eecs.qmul.ac.uk
s.abdallah@ucl.ac.uk

ABSTRACT

The Melody Triangle is a smartphone application for Android that lets users easily create musical patterns and textures. The user creates melodies by specifying positions within a triangle, and these positions correspond to the information theoretic properties of generated musical sequences. A model of human expectation and surprise in the perception of music, *information dynamics*, is used to 'map out' a musical generative system's parameter space, in this case Markov chains. This enables a user to explore the possibilities afforded by Markov chains, not by directly selecting their parameters, but by specifying the subjective *predictability* of the output sequence. As users of the app find melodies and patterns they like, they are encouraged to press a 'like' button, where their setting are uploaded to our servers for analysis. Collecting the 'liked' settings of many users worldwide will allow us to elicit trends and commonalities in aesthetic preferences across users of the app, and to investigate how these might relate to the information-dynamic model of human expectation and surprise. We outline some of the relevant ideas from information dynamics and how the Melody Triangle is defined in terms of these. We then describe the Melody Triangle mobile application, how it is being used to collect research data and how the collected data will be evaluated.

1. INTRODUCTION

The use of generative stochastic processes in music composition has been widespread for decades—for instance Iannis Xenakis applied probabilistic mathematical models to the creation of musical materials [1]. However it can sometimes be difficult for a composer to find desirable parameters and navigate the possibilities of a generative algorithm intuitively.

The Melody Triangle is an interface for the discovery of melodic content where the parameter space of a stochastic generative musical process, the Markov chain, is 'mapped out' according to the *predictability* of the output. The Melody Triangle was developed in the context of *information dynamics* [15]; an information theoretic approach to modelling human expectation and surprise in the perception of music. Users of the Melody Triangle do not select the parameters to generative processes directly, rather they provide input in the form of a position within a triangle, and this maps to the information theoretic properties of an output melody. For instance one corner of the triangle returns completely random melodies, while an other area yields entirely predictable and periodic patterns, the entirety of the triangle covering a spectrum of predictability of the output melodies.

In section 2 we review the concepts and ideas behind information dynamics, and outline the information measures that lead to the development of the Melody Triangle, which have been described in greater detail in our previous work [15]. In section 3 we describe how these information measures are used to construct the Melody Triangle, and how the triangular interface is used to retrieve patterns of symbols that are then mapped to notes or percussive sounds. The Melody Triangle has in previous work been implemented as an interactive installation and as a desktop application, these implementations are described and evaluated in [19]. In section 4 we describe the Melody Triangle mobile app for Android, which is the main contribution of this paper. We outline its features, how it allows users to share their settings with each other, and how it is currently being used to collect data for research. We then describe how the collected data will be interpreted to identify trends and commonalities in aesthetic preferences across users of the app, and to determine if parallels between these preferences and the information dynamics models can be made.

2. INFORMATION DYNAMICS

The relationship between Shannon's [3] information theory and music and art in general has been the subject of some interest since the 1950s [4–8]. The general thesis is that perceptible qualities and subjective states like uncertainty, surprise, complexity, tension, and interestingness are closely related to information-theoretic quantities like entropy, relative entropy, and mutual information.

Music is an inherently dynamic process. An essential aspect of this is that music is experienced as a phenomenon that unfolds in time, rather than being apprehended as a static object presented in its entirety. Meyer [9] and Narmour [10] argued that the experience depends on how we change and revise our conceptions *as events happen*, on how expectation and prediction interact with occurrence, and that, to a large degree, the way to understand the effect of music is to focus on this 'kinetics' of expectation and surprise.

Prediction and expectation are essentially probabilistic concepts and can be treated mathematically using probability theory. We suppose that when we listen to music, expectations are created on the basis of our familiarity with various styles of music and our ability to detect and learn statistical regularities in the music as they emerge. There is experimental evidence that human listeners are able to internalise statistical knowledge about musical structure, [11], and also that statistical models can form an effective basis for computational analysis of music, [12–14].

Information dynamics considers several different kinds of predictability in musical patterns, how these might be quantified using the tools of information theory, and how they shape or affect the listening experience. Our working hypothesis is that listeners maintain a dynamically evolving probabilistic belief state that enables them to make predictions about how a piece of music will continue.

They do this using both the immediate context of the piece as well as using previous musical experience, such as a familiarity with musical styles and conventions. As the music unfolds, listeners continually revise this belief state, which includes predictive distributions over possible future events. These changes in probabilistic beliefs can be associated with quantities of information; these are the focus of information dynamics.

In this next section we briefly describe the information measures that we use to define the Melody Triangle, however a more complete overview of information dynamics and some of its applications can be found in [15] and [2].

2.1 Sequential Information Measures

Consider a sequence of symbols from the viewpoint of an observer at a certain time, and split the sequence into a single symbol in the *present* (X_t), an infinite *past* ($\overleftarrow{X}_t$) and the infinite *future* ($\overrightarrow{X}_t$). The symbols arrive at a constant, uniform rate.

The *entropy rate* of a random process is a well-known, basic measure of its randomness or unpredictablity. The entropy rate is the entropy, *H*, of the *present* given the *past*:

$$h_\mu = H(X_t|\overleftarrow{X}_t). \quad (1)$$

that is, it represents our average uncertainty about the present symbol *given* that we have observed everything before it. Processes with zero entropy rate can be predicted perfectly given enough of the preceding context.

The *multi-information rate* ρ_μ [16] is the mutual information, *I*, between the 'past' and the 'present':

$$\rho_\mu = I(\overleftarrow{X}_t; X_t) = H(X_t) - H(X_t|\overleftarrow{X}_t). \quad (2)$$

Multi-information rate can be thought of as measures of *redundancy*, quantifying the extent to which the same information is to be found in all parts of the sequence. It is a measure of how much the predictability of the process depends on knowing the preceding context. It is the difference between the entropy of a single element of the sequence in isolation (imagine choosing a note from a musical score at random with your eyes closed and then trying to guess the note) and its entropy after taking into account the preceding context: If the previous symbols reduce our uncertainty about the present symbol a great deal, then the redundancy is high. For example, if we know that a sequence consists of a repeating cycle such as . . . b, c, d, a, b, c, d, a . . . , but we don't know which was the first symbol, then the redundancy is high, as $H(X_t)$ is high (because we have no idea about the present symbol in isolation), but $H(X_t|\overleftarrow{X}_t)$ is zero, because knowing the previous symbol immediately tells us what the present symbol is.

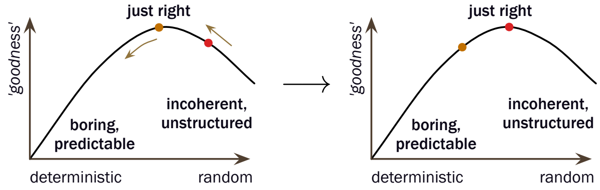

Figure 1: The Wundt curve relating randomness/complexity with perceived value. Repeated exposure sometimes results in a move to the left along the curve [18].

The *predictive information rate* (PIR) [15] brings in our uncertainty about the future. It is a measure of how much each symbol reduces our uncertainty about the future as it is observed, *given* that we have observed the past:

$$b_\mu = I(X_t; \overrightarrow{X}_t|\overleftarrow{X}_t) = H(\overrightarrow{X}_t|\overleftarrow{X}_t) - H(\overrightarrow{X}_t|X_t, \overleftarrow{X}_t). \quad (3)$$

It is a measure of the mutual information between the 'present and the 'future given the 'past'. In other words, it is a measure of the *new* information in each symbol.

The behaviour of the predictive information rate make it interesting from a compositional point of view. The definition of the PIR is such that it is low both for extremely regular processes, such as constant or periodic sequences, *and* low for extremely random processes, where each symbol is chosen independently of the others, in a kind of 'white noise'. In the former case, the pattern, once established, is completely predictable and therefore there is no *new* information in subsequent observations. In the latter case, the randomness and independence of all elements of the sequence means that, though potentially surprising, each observation carries no information about the ones to come. Processes with high PIR maintain a certain kind of balance between predictability and unpredictability in such a way that the observer must continually pay attention to each new observation as it occurs in order to make the best possible predictions about the evolution of the sequence. This balance between predictability and unpredictability is reminiscent of the inverted 'U' shape of the Wundt curve (see Fig. 1), which summarises the observations of Wundt [17] that stimuli are most pleasing at intermediate levels of novelty or disorder, where there is a balance between 'order' and 'chaos'.

A similar shape is visible in the upper envelope of the

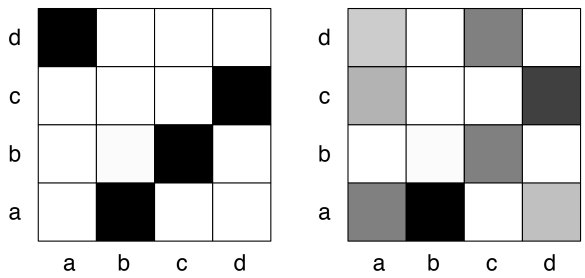

Figure 2: Two transition matrixes representing Markov chains. The shade of grey represents the probabilities of transition from one symbol to the next (white=0, black=1). The current symbol is along the bottom, and the next symbol is along the left. The left hand matrix has no uncertainty; it represents a periodic pattern (a,d,c,b,a,d,c,b,a,d,c,b,a…). The right hand matrix contains unpredictability but nonetheless is not completely without perceivable structure (we know for instance that any 'b' will always be followed by an 'a' and preceded by a 'c'), it is of a higher entropy rate.

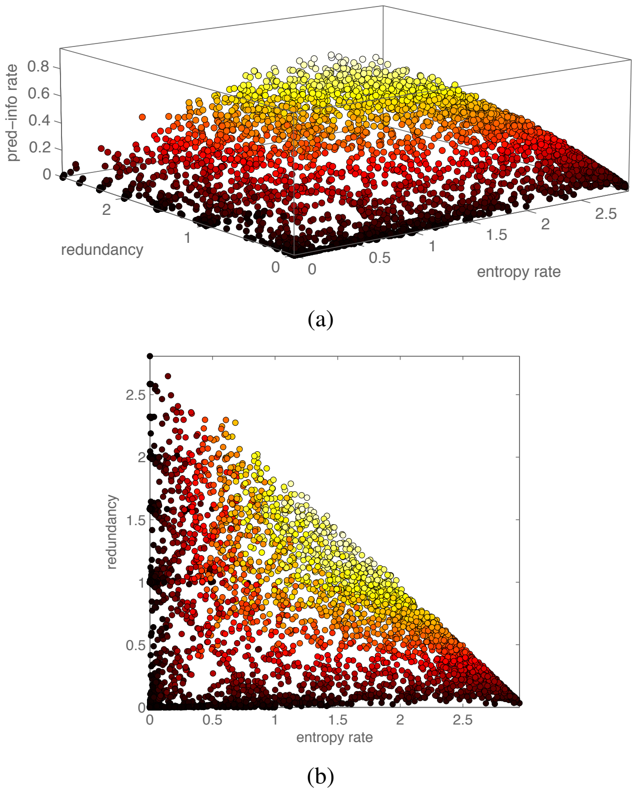

Figure 3: The population of hundreds of randomly generated 8-state transition matrices in the 3D space of entropy rate (h_μ), redundancy (ρ_μ) and predictive information rate (b_μ), all in bits. As can be seen in (a) the distribution as a whole makes a curved sheet, with the highest PIR values found at intermediate entropy and redundancy. Although not visible in this plot, it is largely hollow in the middle. As can be seen in (b), the same plot with the PIR dimension projected out forms a right angled triangle, this is the triangle which corresponds to the interface of the Melody Triangle.

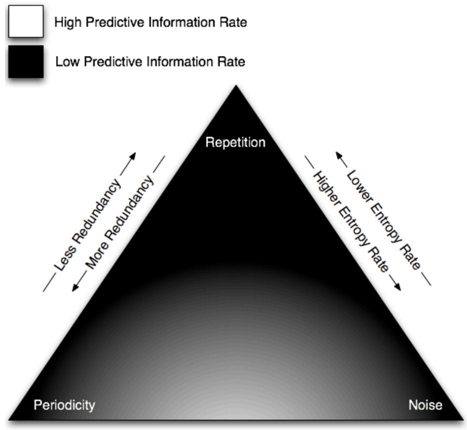

Figure 4: The Melody Triangle

plot in Fig. 3a, which is a 3-D scatter plot of the information measures for several thousand first-order, eight state Markov chain transition matrices generated by a random sampling method. The coordinates of the 'information space' are entropy rate (h_μ), redundancy (ρ_μ), and predictive information rate (b_μ). The points along the 'redundancy' axis correspond to periodic Markov chains. Those along the 'entropy' axis produce uncorrelated sequences with no temporal structure. Processes with high PIR are to be found at intermediate levels of entropy and redundancy.

These observations led us to construct the 'Melody Triangle'.

3. THE MELODY TRIANGLE

The Melody Triangle is an interface that is designed around this natural distribution of Markov chain transition matrices in the information space of entropy rate (h_μ), redundancy (ρ_μ) and predictive information rate (b_μ), as illustrated in Fig. 3a.

The distribution of transition matrices in this space forms a relatively thin curved sheet. Thus, it is a reasonable simplification to project out the third dimension (the PIR) and present an interface that is just two dimensional, resulting in a right-angled triangle, as can be seen in Fig. 3b.

The right-angled triangle is rotated, reflected and stretched to form an equilateral triangle with the 'redundancy'/'entropy rate' vertex at the top, the 'redundancy' axis down the left-hand side, and the 'entropy rate' axis down the right, as shown in Fig. 4. This is our 'Melody Triangle' and forms the interface by which the system is controlled.

3.1 Usage

The user selects a point within the triangle, this is mapped into the information space and the nearest transition matrix is used to generate a sequence of values which are then sonified either as pitched notes or percussive sounds.

Though the interface is 2D, the third dimension (predictive information rate) is implicitly present, as transition matrices retrieved from along the centre line of the triangle will tend to have higher PIR. As shown in Fig. 4, the corners correspond to three different extremes of predictability and unpredictability, which could be loosely characterised as 'periodicity', 'noise' and 'repetition'. Melodies from the 'noise' corner (high h_μ, low ρ_μ and low b_μ) have no discernible pattern; those along the 'periodicity' to 'repetition' edge are all cyclic patterns that get shorter as we approach the 'repetition' corner, until each is just one repeating note. Those along the opposite edge consist of independent random notes from non-uniform distributions. Areas between the left and right edges will tend to have higher predictive information rate (b_μ), and we hypothesise that, under the appropriate conditions, these will be perceived as more 'interesting' or 'melodic.' These melodies have some level of unpredictability, but are not completely random. Or, conversely, are predictable, but not entirely so.

Given coordinates corresponding to a point in the triangle, we select from a pre-built library of random processes, choosing one whose entropy rate and redundancy match the desired values. The implementations discussed in this paper use first order Markov chains as the content generator, since it is easy to compute the theoretically exact values of entropy rate, redundancy and predictive information rate given the transition matrix of the Markov chain. However, in principle, any generative system could be used to create the library of sequences, given an appropriate probabilistic listener model supporting the estimation of entropy rate and redundancy.

The Markov chain based implementation generates streams of symbols in the abstract; the alphabet of symbols is then mapped to a set of distinct sounds, such as pitched notes in a scale or a set of percussive sounds. By layering these streams, intricate musical textures can be created. The number of states in the generated Markov chains corresponds to the number of audio samples used, however the output of the Melody Triangle could even be mapped to non sonic outputs such as visible shapes, colours, or movements.

The information measures that define the Melody Triangle assume a constant rate of symbols, and thus the output sequences proceed at a constant, uniform rate. Although the placing of events in time and rhythm has a strong effect on expectations, surprise and satisfaction in music, the system does not, as yet, address this temporal dimension. Additionally the system does not address the culturally defined expectations of melodic structure that result from our exposure to tonal music; all symbols are considered equal, regardless of what note in a scale they are mapped to.

4. THE MOBILE APP

The Melody Triangle has been implemented as an interactive multi-user installation, as a desktop composition tool, and most recently as a mobile app for the Android platform. It was launched on 28th March 2013, and is free to download from the Google Play app store. [1] A description of the interactive installation and the desktop versions of the Melody Triangle, as well as some user trials can be read in [19].

[1] The download link and some sample audio can be found at http://melodytriangle.eecs.qmul.ac.uk/

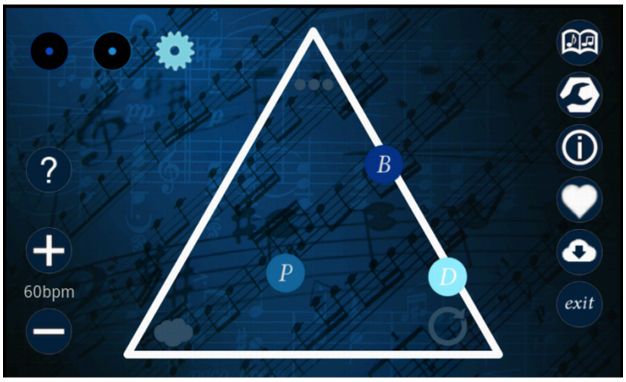

Figure 5: Screenshot of the Melody Triangle mobile app for Android. The letters on the tokens correspond to the instrument they are currently assigned to. P=piano, B=bass, D=drums.

To support the crowdsourcing of data, the app needs to provide enough musical variety to engage users. A simple implementation (with for instance, one single concurrent melody, at one single rate and timbre), would make data analysis easier and more straight forward, however the limit musical appeal would make it difficult to collect data from the public. In the next sections we outline the features of the app, describe how the data is collected, and how it will be analysed.

4.1 Features

As seen in Fig. 5, the app provides three tokens that can be dragged in to the triangle using the touch screen. It is with these tokens that the user selects the points in the triangle that will generate sequences, and thus three sequences can be played simultaneously. Each token can be assigned to one of three instruments: piano, bass, drums. The user can change what instrument is assigned to each token by pressing on the token's holder position on the top left. In addition to changing the instrument, the user can also change the register of the instrument; the piano has three octaves, the bass has two. Additionally the user can select the number of notes per beat, as well as specifying whether this token's notes should be delayed to come on the off-beat, allowing for syncopation between the sequences generated by the tokens.

There are also some global controls; the master beats-per-minute can be changed with the '+/-' buttons on the left, and there is an additional settings menu were the user can choose between the diatonic scale, harmonic scale or the pentatonic scale.

The mobile app is pre-populated with two sets of over 8000 matrixes that densely cover the triangular interface. For the diatonic and harmonic scale (and for the drums samples) the transition matrixes contain 8 states, and for the pentatonic scale 6 states. Whenever a transition matrix is selected by placing a token in the triangle, the symbol-to-note mapping is shuffled. This allows the same transition matrix to correspond to multiple melodies. One state for each of the matrixes is mapped to a rest, allowing for some rhythmic variety and to increase the musicality of the output. When a user taps one of the tokens in the triangle it re-shuffles the symbol-to-note mapping while keeping the same transition matrix.

The current transition matrices, settings of each token and the global settings constitute the 'state' of the system. A user can save their favourites states locally as presets or share them with the world by pressing the 'like' button.

4.2 Collecting Data - 'Likes' and the Melody Triangle 'Radio'

Onscreen notifications encourage users of the app to press the 'like' button (the heart icon on the right of the screen) whenever they enjoy what they are hearing. When they do so, the current state of the system is stored and assigned a unique 6 character hash code, referred to as a 'song id'. The users are given the option to enter a username, or may choose to remain anonymous. This state is encoded into a small file and uploaded to our servers at Queen Mary, University of London. Geographical information is also stored, and a history of previous states is also included in the uploaded information.

Uploaded states become available to other users of the app. By pressing the cloud icon on the right of the screen, the user can type in any song id. When they do so the app downloads the state file from the server and loads the state on to the user's phone. In this way it is possible for users of the app to share settings with each other.

Additionally the app can go into 'radio mode', where the users can quickly and easily audition other users' uploaded states. Upon entering radio mode, the app downloads a randomly selected uploaded state. An additional button appears on the interface, a 'skip' button, which whenever it is pressed the app downloads another randomly selected state. Again the users are encouraged (via on screen notifications) to whenever they enjoy one of the downloaded states, to press the 'like' button. This allows us over time to build a kind of crowdsourced ranking of the uploaded states, as more popular states get more likes. Users can modify downloaded states and then 'like' those, hence states can evolve from other states, and so any uploaded state keeps a history of previous states so that we may track their evolution.

To further encourage uploads and participations, there is a 'Hall of Fame' (see Fig. 6) available at the project website. It shows a list of the users who have contributed the most by uploading many states, as well as chart of most popular songs when 'liked' in radio mode.

In previous work [19] we attempted to carry out a lab study to find links between the information theoretic measures of the Melody Triangle and aesthetic preferences, however it quickly became clear that lab conditions were not practical to get significant amounts of data. The Melody Triangle mobile phone app provides an alternative means of collecting data, while engaging crowds with a unique citizen science project.

4.3 Interpreting Crowdsourced Data

By collecting many liked settings from users all over the world, it may be possible to identify trends and commonalities across these settings. A submitted setting contains all the information relating to the current state of the app, this forms a feature vector that includes the information measures of the currently playing Markov chains, the current note-to-symbol mappings, instrument/register choices, scale, notes per beat for every token and master BPM. Given a submitted state we can extract a number of additional features that are not explicitly stored in the data representing the state of the system, but that are implicitly available by observing the output. This includes the frequencies of notes and melodic intervals for each melody, and by looking across concurrent melodies, inter-melody intervals allowing us to extract harmonic information.

We can look for clusters in the feature space to answer a variety of questions. For instance we can identify what the most common intervals are, both within a melody, and across concurrent melodies, and whether these correspond to the more consonant intervals. We can look for the average information values of the Markov chains, and see how these vary based on the number of concurrent tokens, the rate at which notes are output, or register for instance. We can see if the states that receive the greatest like-to-download ratio in 'Radio mode' have similar information properties to each other.

We are in the active state of research [2] and a full analysis is yet to be carried out. However it is already clear from data collected so far that the more 'predictable' half of the triangle (the half with lower entropy rate and higher redundancy) is preferred to the 'unpredictable' half of the triangle. Additionally it has been observed that the visual layout of the interface has an influence on the parameter choices; a number of states contain tokens lined up in rows or columns. Approximately 20% of states submitted so far contain only the drum sounds, and these may lend themselves to a more straight-forward information theo-

TOP UPLOADERS

Username	Number of Uploads
Anonymous users	96
BLUE	15
MRSPACEMAN	12
FFFFRED	9
HDC	7
EX	7
CONKI	4

TOP SONGS

Chart Position	SongID	Composer	Recent Likes
1	T31FFJV	EX	7
2	MW9PCSG	EX	6
3	TOEM3G8	*Anonymous user*	5
3	TMREXLE	*Anonymous user*	5
3	M7GEMIO	*Anonymous user*	5
3	TWQNOXT	*Anonymous user*	5

Figure 6: The Melody Triangle 'Hall of Fame' as of 9th of June 2013. The top list shows the most prolific users who have shared the most settings by pressing the 'like' button. The lower list shows the top ranked songs based on the number of 'likes' a state has received by other users while in 'Radio Mode'. The hall of fame can be found at http://melodytriangle.eecs.qmul.ac.uk/

[2] As of June 9th 2013, there have been 173 submitted settings

retic analysis as these are not subject to cultural melodic expectations.

Clusterings in the state-space of the data may provide us with the means to link the information dynamic models and its measures to aesthetic preferences. Additionally if we get enough entries, the geographical information may allow us to determine if there are any cultural differences between users based on countries or continents.

5. CONCLUSION

We presented the Melody Triangle; an interface for the discovery of melodic content where the input — positions within a triangle — corresponds to the predictability of the output melodies. The Melody Triangle is contextualised in *information dynamics*; an information theoretic approach to modelling human expectation and surprise. We outlined the relevant ideas behind information dynamics and described three key information theoretic measures; entropy rate, redundancy and a measure of *predictive information rate*, which describes the gain in information made by current observations about the future, but which are not already known from past observations. We described how the natural distribution of randomly generated Markov chains in terms of these measures lead us to design the Melody Triangle.

We described the Melody Triangle mobile app, a free app for Android, and outlined how it collects data for research by uploading the 'liked' settings of users to our servers. We describe the app's 'radio mode' that enables users to quickly audition other uploaded states provide feedback to form a crowd-sourced rankings table of most popular settings. Finally we outline how the collected data will be used to look for trends and commonalities in the uploaded settings, and to help identify any relationship between the information-dynamic model of human expectation and aesthetic preference.

Acknowledgments

This work is supported by an EPSRC Doctoral Training Centre EP/G03723X/1 (HE), GR/S82213/01 and EP/E045235/1(SA), an EPSRC Leadership Fellowship, EP/G007144/1 (MDP) and EPSRC IDyOM2 EP/H013059/1. The Melody Triangle mobile app was developed with QApps and supported by impactQM, funded by the EPSRC.

6. REFERENCES

[1] I. Xenakis, *Formalized music : thought and mathematics in composition*. Stuyvesant, NY: Pendragon Press, 1992.

[2] S. A. Abdallah, H. Ekeus, P. Foster, A. Robertson, and M. D. Plumbley, "Cognitive music modelling: An information dynamics approach," *Cognitive Information Processing (CIP), 2012 3rd International Workshop on*, pp. 1–8, 2012.

[3] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423,623–656, 1948.

[4] J. E. Youngblood, "Style as information," *Journal of Music Theory*, vol. 2, pp. 24–35, 1958.

[5] E. Coons and D. Kraehenbuehl, "Information as a measure of structure in music," *Journal of Music Theory*, vol. 2, no. 2, pp. 127–161, 1958.

[6] A. Moles, *Information Theory and Esthetic Perception*. University of Illinois Press, 1966.

[7] L. B. Meyer, *Music, the arts and ideas: Patterns and Predictions in Twentieth-century culture*. University of Chicago Press, 1967.

[8] J. E. Cohen, "Information theory and music," *Behavioral Science*, vol. 7, no. 2, pp. 137–163, 1962.

[9] L. Meyer, "Music, the arts, and ideas: Patterns and Predictions in Twentieth-Century Culture," University of Chicago Press, Chicago, 1967.

[10] E. Narmour, *Beyond Schenkerism*, ser. the need for alternatives in music analysis. Univ of Chicago Press, 1977.

[11] J. R. Saffran, E. K. Johnson, R. N. Aslin, and E. L. Newport, "Statistical learning of tone sequences by human infants and adults," *Cognition*, vol. 70, no. 1, pp. 27–52, 1999.

[12] D. Conklin and I. H. Witten, "Multiple viewpoint systems for music prediction," *Journal of New Music Research*, vol. 24, no. 1, pp. 51–73, 1995.

[13] D. Ponsford, G. A. Wiggins, and C. S. Mellish, "Statistical learning of harmonic movement," *Journal of New Music Research*, vol. 28, no. 2, pp. 150–177, 1999, also available as Research Paper 874, from the Division of Informatics, University of Edinburgh.

[14] M. T. Pearce, "The construction and evaluation of statistical models of melodic structure in music perception and composition," Ph.D. dissertation, Department of Computing, City University, London, 2005.

[15] S. Abdallah and M. Plumbley, "Information dynamics: patterns of expectation and surprise in the perception of music," *Connection Science*, vol. 21, no. 2, pp. 89–117, 2009.

[16] S. Dubnov, "Generalization of spectral flatness measure for non-gaussian linear processes," *Signal Processing Letters, IEEE*, vol. 11, no. 8, pp. 698–701, 2004.

[17] W. Wundt, *Outlines of Psychology*. Lepzig: Englemann, 1897.

[18] D. E. Berlyne, *Aesthetics and Psychobiology*. New York: Appleton Century Crofts, 1971.

[19] H. Ekeus, S. Abdallah, and M. Plumbley, "The Melody Triangle: Exploring Pattern and Predictability in Music," *Musical Metacreation (MUME), 1st International Workshop on*, 2012.