Analysis of local and global timing and pitch change in ordinary melodies

Roger Watt
Dept. of Psychology, University of Stirling, Scotland
r.j.watt@stirling.ac.uk

Sandra Quinn
Dept. of Psychology, University of Stirling, Scotland
s.c.m.quinn@stirling.ac.uk

ABSTRACT

This paper describes a set of statistical relationships between pitch change structure and timing structures in ordinary melodies. We obtained over 5000 MIDI files for ordinary western melodies, each with a prescribed tempo, so that note timings could be given in seconds. 1): We find that the frequencies of occurrence of different pitch change sizes are stationary: they do not vary during the time-course of a melody, apart from during the first 1 second and the final 1 second. 2): There is an inverse relationship between the mean (absolute) pitch change size in a melody and the mean time interval between successive note onsets: melodies with larger pitch changes tend to be faster.3): The time intervals between successive occurrences of the same pitch change size reflect an active process. 4): For each melody, we construct a function showing the temporal rise and fall in the likelihood of the melody as given by the log of the reciprocal of the frequency of the most recent pitch change. Fourier analysis of these functions shows a regular pattern of coherent variability with a period of between 2 and 6 seconds. Low likelihood portions of a melody are balanced by higher likelihood ones over a time scale of a few seconds.

Keywords

Melody Timing Pitch Statistics

INTRODUCTION

It is widely recognized that Western music operates on the basis of the combined effects of pitch changes and local timing differences. A melody can have a profoundly different effect if either pitch structure or timing structure is changed. The purpose of the studies reported in this paper is to explore whether there are any robust statistical patterns which relate pitch structure and temporal structure.

There has been some considerable statistical research into pitch structures in various genre of music. More recently, several large datasets have been obtained for statistical analysis, such as the Essen dataset (Schaffrath, 1992; Selfridge-Field, 1995). Datasets of this type have been put to various purposes, broadly musicological in nature, such as establishing the frequency of melodic arches in folksong (Huron, 1996); the relationship between the nature of melodic features and the source location for the melody (Aarden and Huron, 2001).

The present studies focus on the temporal properties of melodies. Rather than using a specific delimited dataset of melodies, we have used a rather broad range attempting to capture the variability in what might be called ordinary musical experience. Our purpose here is not musicological, per se, but is essentially perceptual. When a non-trained listener hears an ordinary melody, they are undoubtedly processing it with some sophistication – even if their lack of training precludes them from describing the musical nature of the experience. Our fundamental proposition is twofold: first we propose that listeners classify musical events by their frequency of occurrence; and second, that the overall response to a melody is structured fundamentally by the relationship between the unique temporal characters of that melody with respect to temporal regularities they have observed.

So, when a non-trained listener hears a melody start with an upwards 4th, they recognize this as a common event. If a melody were to start with an upwards augmented 4th, they would recognize this as an unusual event. The difference in the effect of these two is then mediated by their respective frequencies of occurrence, not necessarily by any theoretical musical property of the two.

MATERIALS

The research reported uses a collection of over 5,000 melodies taken from MIDI files obtained from a large number of web-sites. All melodies are homophonic throughout. These are a subset of a larger set in excess of 10,000. The set used were selected on the basis that they had exact timing information, so that the note durations in each could be stated in seconds (many MIDI files use a default of 120bpm and it is not clear whether this tempo has any real significance). From the 5000, a random sample of 900 were listened to by two naïve listeners, each of course familiar with the general style, with a view to judging simply whether the melody sounded all-right or not. All were found to be acceptable.

The collection of 5,000 melodies generated a set of over 620,000 notes.

SOFTWARE

All the analysis described below was conducted with software written by the authors in Matlab©. For further details, please contact the first author.

PRELIMINARY ANALYSIS

Before proceeding with the more complex analysis of pitch change and timing, a few simple analyses were performed to further establish the representative nature of the melodies.

Pitch distribution

The distribution of raw pitch across the set of melodies was measured and is shown in Figure 1. A pitch of 60 corresponds to middle C. The distribution is not unexpected, covering the treble staff with a little either side. Figure 2 shows the distribution of pitch changes (from one note to the immediately next note). This is also as would be expected. Note that the term pitch change is used because the more familiar term, interval, could also refer to a time difference.
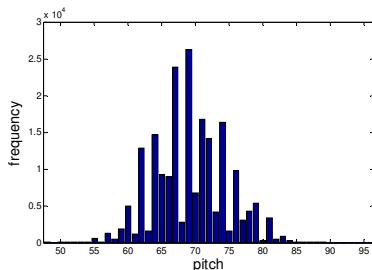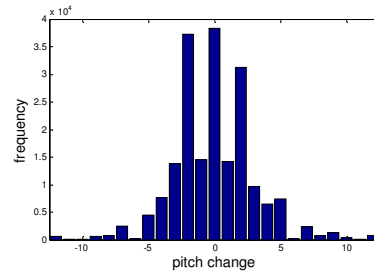

Figure 1: Distribution of pitch

Figure 2: Distribution of pitch changes

Note values durations

Figure 3 shows the distribution of note durations (in seconds as performed). These are also as would be expected, with nearly all under 1 second in duration and more than 2/3 less than half a second.
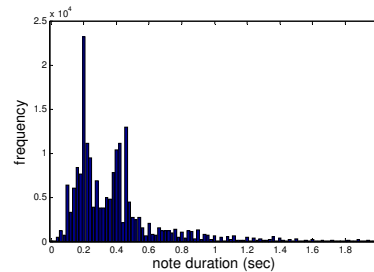

Figure 3: Distribution of note durations

Discussion

This preliminary analysis has demonstrated that the set of melodies have basic structures that are as would be expected for "ordinary melodies" – note pitch mainly within the 2 octaves above middle C; small pitch changes much more frequent than large ones; most note durations less than 0.5 sec.

ANALYSIS 1

The first analysis considers whether the distribution of pitch changes (Δp), as shown in Figure 2, is fixed during the course of a melody, or whether it varies with time. In the extreme limit, it clearly is not: Figure 4 shows the distributions of Δp between the first 2 notes of a melody and between the final 2 notes. As expected, the distributions are more selective.
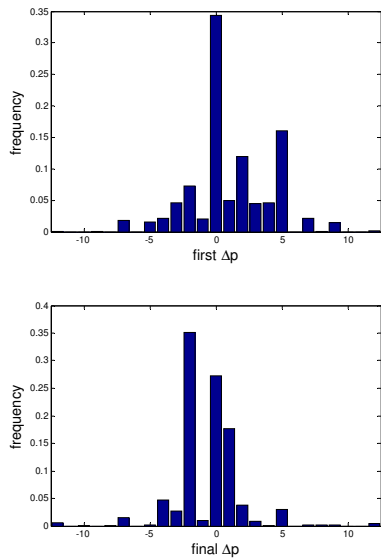
Figure 4: Distributions of first and final pitch changes

The issue is whether the instances of any particular pitch change size tend to happen more in one part of a melody than another. For example, a Δp of size +6 is uncommon. Do the few instances of that pitch change size tend to occur late in a melody or not?

A time base was constructed with a duration of 20 seconds, and a resolution of 0.1 second. Starting with a particular Δp, say +4 for example, all the instances of this event were placed at the appropriate place on this time base, producing a function showing how the probability of that Δp occurring changes over time. This function is shown in Figure 5 for several different values of Δp. As can be seen, after the first 2 seconds there is little or no variation in frequency: the process is stationary.
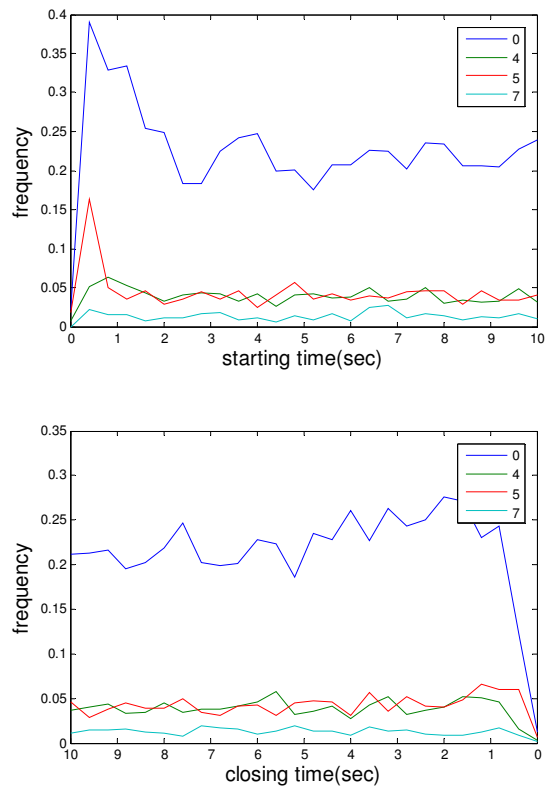

Figure 5: variations in frequency of Δp with time. Data are shown for Δp = 0, +4, +5, +7

The resultant pattern was analysed by piecewise linear least squares fit to establish whether there was any significant trend between the occurrence of pitch changes and time. The finding is that, excluding the first 2 seconds and the final 2 seconds, there is no trend for any of the different Δp values.

Figure 6 shows how the magnitude of this non-stationarity for the first opening and ending 2 seconds as a function of the pitch change size. The figure also shows the same data, but plotted against the frequency of occurrence pitch change. The latter function is simple in form and shows a broadly linear effect.
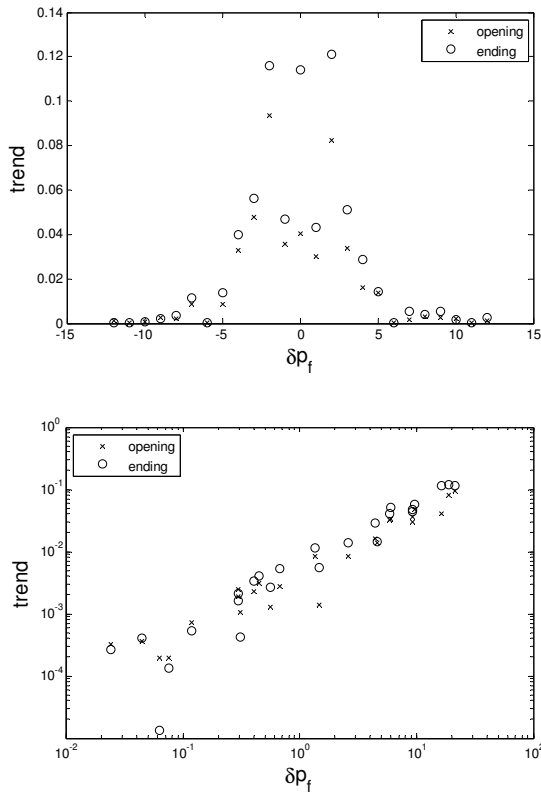
Figure 6: The size of the change in frequency over the opening (x) and closing (o) 2 seconds of melodies, plotted as a function of Δp (top) and frequency of occurrence of pitch change (bottom)

ANALYSIS 2

The second main analysis concerns the relationship between pitch change size and the durations of the notes immediately either side of them. The first issue to mention is that pitch changes in a melody do not have a duration: it is the notes either side of a pitch change that have a duration. In the data that follow, we will use negative values to denote the duration of the note that precedes a pitch change, and positive values for the note that follows the pitch change.

The first set of data, shown in Figure 7, explore whether there is any relationship between note durations and Δp. As can be seen, there is a tendency for pitch changes greater than 4 semitones to be both preceded and (to a lesser extent) followed by longer notes. That longer notes tend to be followed by a larger pitch change (left side of top panel) is not surprising: in most simple melodies, a long note is often a phrase boundary. The smaller tendency for larger pitch changes to be followed by a longer duration note (right side of lower panel) is more interesting.
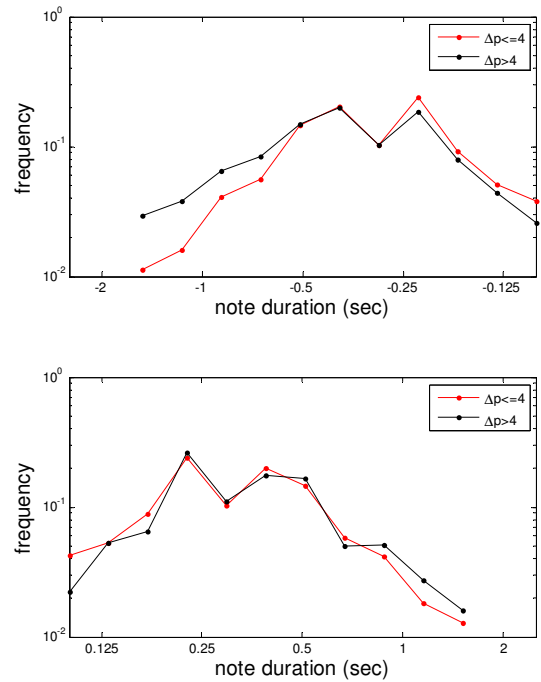

Figure 7: Frequency of note duration and Δp combinations

For any given melody we can calculate the mean note duration, which is a measure of how fast the melody is played (ie the inverse of its tempo in notes per second), and we can calculate the mean absolute pitch change size (absolute Δp is the unsigned value). We are interested to establish whether these two properties might be related. Data above show a weak relationship between individual Δp and note duration, and maybe a melody with higher than average pitch change sizes will have a slightly different tempo.

The next analysis concerns the relationship between the mean (absolute) Δp and the mean note duration, for each melody. Figure 8 shows a function relating these two. The horizontal scale is the mean pitch change. For each plotted value, all melodies with the appropriate mean pitch changes were taken, and the mean note duration for that set of melodies was then calculated. The vertical scale is seconds. It can be seen that there is a broad trend visible, with increased mean Δp values tending to correspond with lower mean note durations.
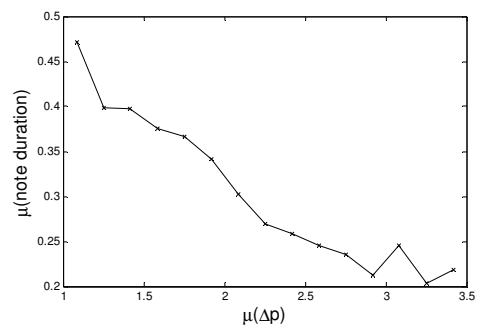

Figure 8: Mean note duration as a function of mean Δp

ANALYSIS 3

We turn now to explore some of the broader temporal structure in the melodies. The earlier analysis of time of occurrence of Δp events showed that, on the average across melodies, the likelihood of any particular Δp occurring does not vary. This could mean two things. First, the same could be true for an individual melody because there is no temporal structure in melodies. Alternatively, the occurrence of a particular Δp could be highly structured within a melody, but when averaged across melodies, that structure is hidden. For example, suppose that after each occurrence of a Δp with value -4, there is a period of 3 notes following during which that Δp will never occur, then there is an important structure within a melody that will be averaged out across the population if the first occurrence occurs randomly. We explore this possibility now.

Figure 9 shows three distributions. Each is the distribution of waits from one occurrence of a given Δp to the next. The waiting time is normalized to the average wait for each (so all three distributions have the same mean of 1). As can be seen the shapes of the distributions are rather different.
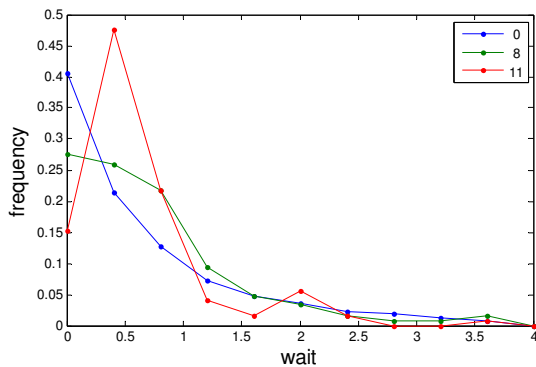

Figure 9: Distributions of waits for 3 different pitch change size

If the timings of a sequence of events (such as the occurrences of $\Delta p = -4$) are random and completely independent of each other, then the distribution time of intervals between successive occurrences will follow the exponential distribution. The exponential distribution has only one parameter, the mean interval between occurrences. The exponential distribution is a special case of the gamma distribution, which has a second parameter. The second parameter in the gamma distribution is independent of the first parameter and relates to the shape of the distribution. It can be characterized as the index of the earliest arrival of the next occurrence of interest in the event stream. If the gamma shape parameter equals 1 (as in the exponential case), then the next occurrence of a $\Delta p=-4$ could be the very next event. If the gamma shape parameter equals 3, then the next occurrence will definitely not happen in the next 2 events. So in a melody, if there is a silent inhibitory period after a $\Delta p=-4$ of 3 events (note that the measure of

this period is in events, not in seconds of actual time), then the distribution of time intervals (in actual seconds, not event counts!) between such occurrences will be a gamma distribution with the shape parameter set to 4.
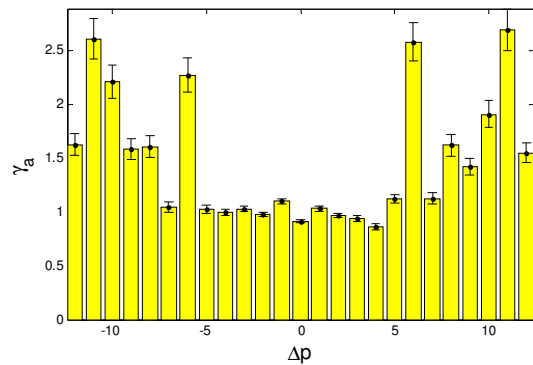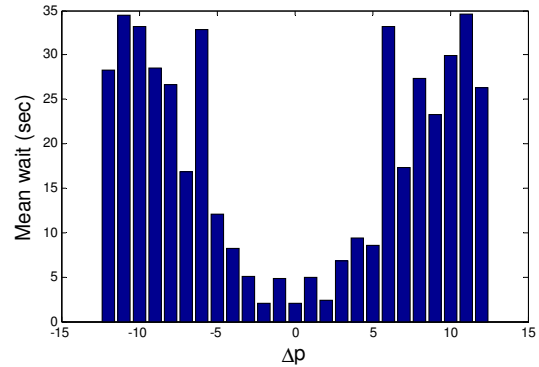

Figure 10: (Top) The mean time to wait between successive occurrences of different Δp
(Bottom) Gamma shape parameter as a function of Δp

We have collected together all the time intervals between successive occurrences for each Δp value in turn. The gamma distribution parameters for these distributions of time intervals can be estimated and are shown in Figure 10. The top graph shows the mean time interval, and as would be expected this is larger for the less common Δp values. The lower graph shows the gamma shape parameter γ_a. For smaller values of Δp between 0 and 5, plus 7, this is close to 1, indicating that the occurrence of one of these events does not inhibit a second one: in this statistical sense they are a random process. For larger values of Δp the gamma shape parameter is much higher – indicating non-random statistical structure.

Figure 11 shows the gamma shape parameter as a function of the frequency of occurrence of the pitch change. As can be seen, this graph has a simple form, with a linear effect for less common pitch changes (with a frequency of occurrence less than 0.01) and little or no effect for the more common pitch changes.
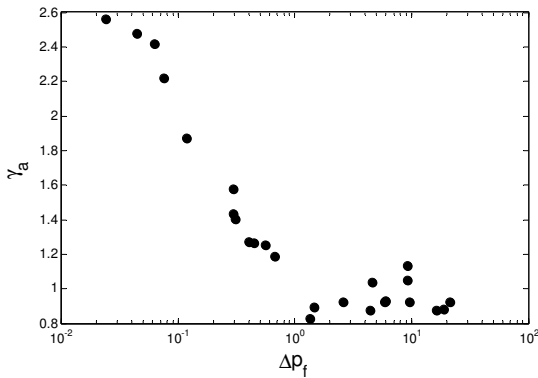
Figure 11: gamma shape parameter as a function of the frequency of occurrence of the Δp

ANALYSIS 4

The final analysis explores this issue of temporal scale further.

We can start with one of the melodies. This is a sequence of pitch changes, each one with a specific moment in time. Figure 12 shows an example: at the top is a representation of the pitch structure in time; beneath this is a representation of the pitch changes as a function of time. Pitch changes are effectively instantaneous and for the duration of each note the pitch change function is set to zero.

For each pitch change we can calculate an a priori probability simply from the distribution of Δp across all melodies. For rare pitch changes, this value will be very small. Figure 13 (top) shows the function obtained when pitch change is replaced by this value. The resultant function clearly has a pattern of regular changes with fairly well defined durations. For example, there is a pattern with a period of around 5 seconds which is shown in the bottom of Figure 13. This pattern has quite a large energy (the changes in frequency of occurrence are quite marked).

We can calculate in this way the variation in energy for all possible periods. This pattern has two main contributions: the temporal pattern of note onsets themselves and the temporal pattern of Δp. We can remove the effect of the note onsets, and then we are left with a function, shown in the bottom row of Figure 14 which shows the amount of energy as a function of period (called a power spectrum). This graph shows that there is a large amount of energy for periods of around 2 seconds and again for periods between 4 and 8 seconds. These temporal variations in the melody are caused by the pattern of use of pitch changes.
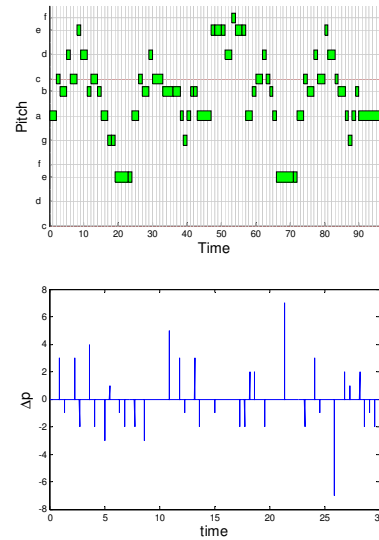

Figure 12: a melody as a temporal pitch structure (above) and as a temporal pitch change function (below)
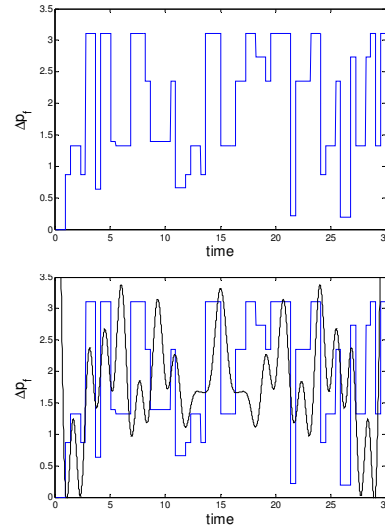

Figure 13: The same melody represented as frequency of occurrence of the most recent pitch change as a function of time (top) with one of the strong periodic patterns (bottom)
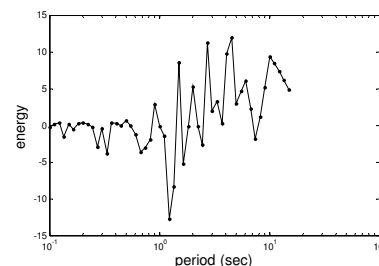

Figure 14: the power spectrum for the temporal variations in frequency of occurrence

Figure 15 shows the average power spectrum for the tempo-ral variation in the frequency of occurrence of pitch changes for all melodies, with confidence limits. As can be seen there is a very marked increase in energy for periods of between 2 and about 6 seconds. On this analysis, there is no evidence for structure at longer periods.
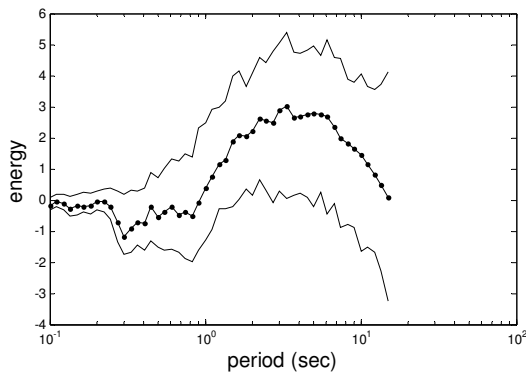

Figure 15: Power spectrum for the variation in event fre-quency of occurrence over time in melodies

SUMMARY OF RESULTS

This paper presents the results of 4 analyses of the temporal structure of a large set of ordinary melodies. These melo-dies are all drawn from, and are ordinary in the sense that they are commonplace. The findings relate therefore only to such melodies.

The analysis has focused on the temporal properties of the various different pitch changes found in such melodies. The use of pitch change has an obvious practical benefit in that it avoids the complications associated musical key that would arise if pitch per se were used. However, it is also important because pitch changes are already a highly local-ized event in time.

The different analyses all concerned temporal properties at different time scales from the instantaneous through to sev-eral seconds.

The first analysis explored whether the distribution of the various different values found for pitch change showed variations in the distribution as a function of time. It was found that the distributions of pitch changes do not vary in time, except for the first 2 seconds and the final 2 seconds of melodies. Over these ranges, the extent to which a pitch change varies in its probability of occurrence is a simple function of its overall probability of occurrences: common pitch changes show a growing probability of occurrence at the starts and ends of melodies. Apart from this latter, case, this analysis is an analysis at an instantaneous time scale.

The second analysis explored the time structure of melo-dies in the local temporal environment of each pitch change event. It was found that the durations of notes either side of a pitch change depended, to some degree on the size of the pitch change. For pitch changes greater than 4 semitones, the preceding and the following notes will tend to be longer than average.

The same analysis showed an inverse relationship between the mean note duration for a melody and the mean pitch change size. Melodies with more large pitch changes tend to have shorter note durations.

The third analysis considered a longer time scale. The dis-tributions of waits between successive occurrences of a given pitch change size were calculated. This was repeated for each of the various pitch change sizes and the data in each case fit to the gamma distribution. The interest is in the shape parameter of this distribution as this provides a clear indication of the presence of temporal structure. The data show that smaller pitch change sizes tend to occur without restraint in melodies, but that larger ones tend to be spaced wider apart than would be expected on the basis of their frequency of occurrence. This is important because it implies a long range temporal structure in melodies rather than just a strictly local structure. The extent of this longer time scale varies inversely with the frequency of occur-rence of pitch changes: rare pitch changes show the great-est temporal scale effect.

The final analysis considered a different form of temporal structure. We converted each melody into a temporal func-tion showing how common was the most recent event. This function varies with time, and a spectral analysis of the variations was used to establish the typical periodicity of the variations. This periodicity is centred at about 4 sec-onds, and extends a few seconds either side. In other words, the changes from common events to rare events and back again tend to happen with a natural cycle of around 4 seconds. This figure relates to the time scale for melody openings and closings, which was half of this value, but would tend to correspond to half of the cycle.

DISCUSSION

We have identified several different reliable temporal pat-terns with what we are terming ordinary melodies. In dis-cussing this, we will first consider what the musical causes of these patterns might be, and then discuss the psycho-logical and perceptual consequences.

Musically speaking, the data are not unexpected. The tem-poral patterns with a period of 4 second probably corre-spond to short phrases of between 8 and 16 notes. If the starting pitch of one phrase is relatively disconnected from the pitch that ended the preceding phrase, then the lower probability pitch changes will have a higher tendency to occur at phrase boundaries than inside phrases. If more-over, the ends of phrases are sometimes marked by longer than average notes, then this will tend to lead to localized temporal structure around larger pitch changes.

These patterns are best revealed by using a representation of the melody which records the frequency of occurrence of pitch changes rather than pitch changes or pitches them-

selves. This is a useful positive finding from a perceptual and psychological point of view. It suggests that a detailed understanding of musical structure (explicit or implicit) is not required to begin to get a sense of the temporal structure of a melody.

Our process has been, in effect, that which could be conducted by a mind with good statistical competence but little or no musical competence.

ACKNOWLEDGMENTS

REFERENCES
Aarden, B. and Huron, D (2001). Mapping European folk-song: Geographical localization of musical features. In WB Hewlett an E Selfridge Field (Eds) *The Virtual Score: Representation, Retrieval and Restoration* (pp169-183) Cambridge Mass: MIT Press.

Huron, D (1996). The melodic arch in Western folksongs. *Computing in Musicology* 10: 3-23.

Schaffrath (1992). The retrieval of monophonic melodies and their variants: Concepts and strategies for computer-aided analysis. In A Marsden and A Pope (Eds), *Computer Representations and Models in Music* (pp95-109) London: Academic Press.

Selfridge-Field, E (1995). Essen Musical Data Package *(CCARCH Tech Report No. 1)*, Menlo Park Calif.

...