

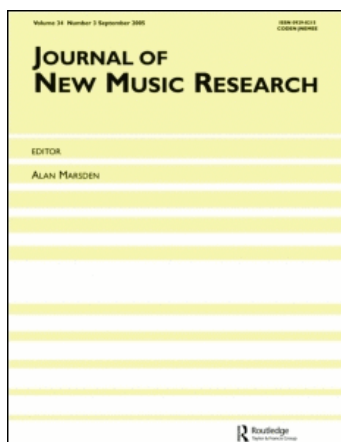
This article was downloaded by: [Aalborg University]

On: 11 May 2011

Access details: Access Details: [subscription number 912902580]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of New Music Research

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713817838>

Recognition of the Main Melody in a Polyphonic Symbolic Score using Perceptual Knowledge

Anders Friberg^a; Sven Ahlbäck^b

^a Speech, Music and Hearing, KTH, Sweden ^b Royal College of Music, Sweden

Online publication date: 30 November 2009

To cite this Article Friberg, Anders and Ahlbäck, Sven(2009) 'Recognition of the Main Melody in a Polyphonic Symbolic Score using Perceptual Knowledge', Journal of New Music Research, 38: 2, 155 – 169

To link to this Article: DOI: 10.1080/09298210903215900

URL: <http://dx.doi.org/10.1080/09298210903215900>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Recognition of the Main Melody in a Polyphonic Symbolic Score using Perceptual Knowledge

Anders Friberg¹ and Sven Ahlbäck²

¹Speech, Music and Hearing, KTH, Sweden; ²Royal College of Music, Sweden

Abstract

It is in many cases easy for a human to identify the main melodic theme when listening to a music example. Melodic properties have been studied in several research projects, however, the differences between properties of the melody and properties of the accompaniment (non-melodic) voices have not been addressed until recently. A set of features relating to basic low-level statistical measures were selected considering general perceptual aspects. A new ‘narrative’ measure was designed intended to capture the amount of new unique material in each voice. The features were applied to a set of scores consisting of about 250 polyphonic ringtones consisting of MIDI versions of contemporary pop songs. All tracks were annotated into categories such as melody and accompaniment. Both multiple regression and support vector machines were applied on either the features directly or on a Gaussian transformation of the features. The resulting models predicted the correct melody in about 90% of the cases using a set of eight features. The results emphasize context as an important factor for determining the main melody. A previous version of the system has been used in a commercial system for modifying ring tones.

1. Introduction

It is in many cases easy for a human to identify the main melodic theme when listening to a given music example. The melodic theme is often considered to be the most characteristic feature of a given musical piece and has

thus been an important topic in many music theory books (e.g. Piston & DeVoto, 1987; Perricone, 2000). Following a more bottom-up approach, statistical analysis of low-level features, such as pitch height, note duration, and pitch interval, has been applied on a large number of melodies using different databases (Huron, 2006). However, finding a more specific description of general properties of what constitutes a melody beyond the low-level statistical properties seems to be a difficult task. Such attempts at specifying melody are rare in the literature. In particular, the difference in specific properties of the melody versus for example the accompaniment voices has until recently not been addressed (Madsen & Widmer, 2007; Rizo et al., 2006).

The purpose of the present investigation is to further investigate the properties related to what we can call *melodicity* of a given part. In particular, we would like to determine the difference between the main melody and the accompaniment. We will start by trying to hypothesize what constitutes a melody in general terms, continue with data from perceptual and statistical studies, followed by a model for predicting the main melody in a polyphonic score. Our starting point was a database of polyphonic ring tone music collected from commercial ring tone providers, where different parts of the music are separated into different MIDI tracks. An early prototype was presented by the authors at the ICMPC conference 2006 in Bologna (Friberg & Ahlbäck, 2006).

2. What is melody?

Melody is a music theoretical concept that is intimately linked to Western music culture. In a general sense,

melody has been used as a concept within Western music theory since classical antiquity and in its current meaning since around 1600. A typical definition of melody can be found in the *Concise Oxford Dictionary of Music* where melody is defined as 'A succession of notes, varying in pitch, which have an organized and recognizable shape.' (Kennedy, 1980). In such a general sense, the concept of melody is relevant within a great number of musical styles around the world, including non-western musics. Thus, a melody is a sequence of notes organized in a way that it 'sticks out' perceptually and is memorable. This definition points to the Gestalt qualities of a note succession as central to the melody concept, and consequently melody was one of the key examples of the Gestalt approach already since the early Gestalt psychologists (von Ehrenfels, 1890). The question is: which *properties* of the tone successions make us perceive it as a melody?

In this context, with regards to finding the melody part in a polyphonic score, the question has to be extended to also ask what it is that makes the melody special in relation to the accompaniment. There are many plausible approaches to this problem. From the point of view of statistical learning, a melody should be similar to other melodies (and dissimilar to the accompaniment) with respect to typical tonal and rhythmical perceptual concepts such as tonal centre, range, interval successions, relation to harmony, metre, etc.

According to Huron (2006, p. 229) the ability of a melody to 'stick' in the ear and be remembered is attributed to repetition. Music is an art form where repetition of structural elements generally plays a crucial role. However, the accompaniment is often the most repeated part in the type of pop music that is featured in the current study and it is still not remembered the same way as the melody. One reason could be that the accompaniment often is representative for a genre or a subgenre rather than specific for each song. Thus, the accompaniment may serve as genre classification (a task that is very quickly determined by an experienced listener) and the melody as the specific selection within the genre. A melody part may then need something that

repeats but not literary and also contrasts with the accompaniment. One common device is to make a successive development of a small musical motive.

The main melody needs also to be heard in the context of both strictly repetitive accompaniment and other auxiliary melodies. For a listener to be able to follow the main melody it should be perceived as one perceptual stream according to stream segregation theory (Bregman, 1990). Thus, there should be some distance to other voices, in terms of timbral contrast, dynamic contrast, melodic interval size, and temporal continuity.

Let's look at some examples in which there may be some difficulties in predicting the main melody. They are all from our music database. The score in Figure 1 shows an example of the main melody and an accompanying melody (the other tracks are not shown). At a first glance one might guess that the top track is the main melody since it has the highest pitch. However, a listener will in most cases recognize the lower track to be the main melody (which is also the melody that is sung in the original version). Why? In this case we can see that the lower track tries to tell us a slightly longer story. In the first measure, a melodic motive is presented that is further developed throughout all the eight bars. The top melody on the other hand contains one motive that is repeated without any variation. Thus the main melody 'tells the story' while the other parts seem to function as to support the main melody.

We now arrive at another possible approach to the concept of melody; melody conceived as a musical narrative, as storytelling by means of pitch and rhythm change. In parallel to how a story constitutes an intelligible whole, composed by series of events that connect and contrast in a dramatically significant way, a melody could be understood as a series of subgestalts—musical phrases—that connects to form a meaningful whole. This means that *melodicity* can be related to the degree to which a successive pitch structure represents the central gestalt level of the piece, the gestalts that 'make up the story' of the piece. Following this reasoning, exact repetition of short substructures would make it harder to perceive the identity of the pitch



Fig. 1. An excerpt of two parts from 'English summer rain' by Placebo.

structure as a whole (see also ‘Measure of narrativity’ below).

An interesting approach is to consider the unique information content in terms of entropy (Madsen & Widmer, 2007). However, since this approach concerns a complexity measure that is not directly related to segmental structure it is not useful for the purpose of the current study.

In the example shown in Figure 2, all the tracks contain short rhythmic figures. This is an instrumental interpretation of a hip-hop song, thus without a sung melody in the original version. The two bars are in fact repeated throughout the whole piece. One might therefore think that there is no main melody in this case. However, listening to the music reveals that the second track can be conceived as the main melody. In the context of the other tracks the second track tells a longer story (two-bar phrase instead of one-bar phrases), thus is the most likely candidate according to the narrative principle referred to in the first example. Thus, this indicates that the concept of main melody might be context dependent relatively to the musical content in the

other tracks, similar to the context dependency of melodic similarity suggested by Ahlbäck (2007).

One might ask if there is a limit to what can be regarded as a melody in terms of change of pitch and duration? Can a repeated note of the same duration be regarded as a melody? The answer to this question is obviously dependent on socio-cultural and individual factors and there are probably people who would not regard the second track in the above example to be a melody. Our approach is, however, relative, which means that we are looking for the track that fits best with the melody concept as it is defined here. For some of the pieces in the database, melody is clearly not the most important, identifying factor, nor the dimension of the greatest structural significance. However, it is interesting to see that producers of these ‘ring tones’ often have focused on the melodic qualities of, for instance, rap songs, which originally have consisted to a lesser degree on what can be considered as melodic material.

The last example in Figure 3 is an ambiguous case in which the melodic focus seems to switch between the two top voices that both have a narrative character. In this case the first voice is introducing a short melodic theme in the first two bars. The second voice is introduced in the third bar. Thus, in this case, the first impression is that the top melody will dominate and that the second melody will serve as a comment to the first.

Melody is also closely related to the concept of song, hence to singing. Typically, a song is in our culture characterized by its melody and is regarded as something fundamentally different from reading or recitation of lyrics. This connection is very obvious in the current study since most of the melodies in the database are actually sung in their original version. Moreover, it is often claimed that a good melody should have vocal qualities, thus easy to sing, implying a certain tessitura, a convenient range, not too wide intervals, and note



Fig. 2. An excerpt from ‘Don’t believe the hype’ by Public Enemy.



Fig. 3. An excerpt from ‘Burn’ by Usher.

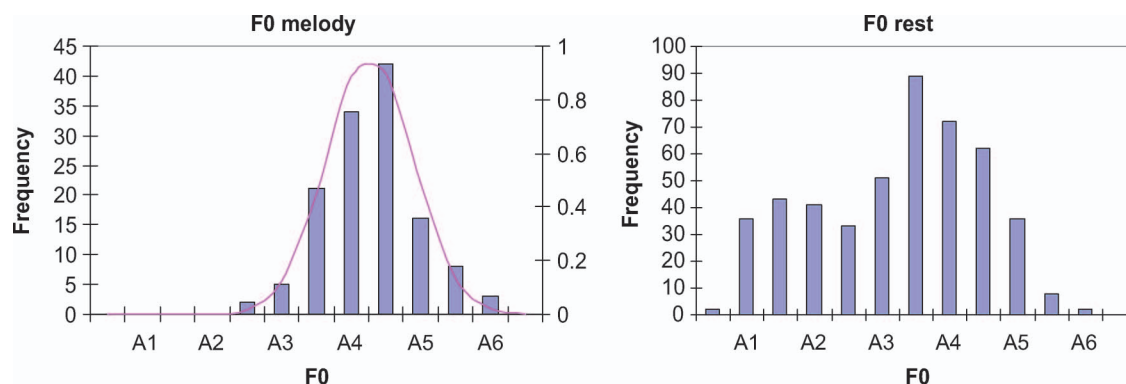


Fig. 4. Histogram of average pitch for the main melody (left) and for all other tracks (right) in the training set.

durations suitable for vocal performance, features that are reflected e.g. in principles for voice leading in Western music.

From the above discussion, it is clear that we are using a high-level feature, the narrative character, as well as a number of low-level features for selecting the main melody. These will be investigated in detail in the next section.

3. Melodic features

We will try to obtain a set of melodic features that may differentiate between the main melody and the accompaniment. The features were selected starting from perceptual/cognitive research along with statistical properties of large collections of melodies, complemented with intuitive knowledge using the principles discussed in the section above. Each feature will be evaluated using the training set of the music database.

3.1 Low-level features

Huron (2001) defined a number of perceptual principles that he ingeniously used for deriving and motivating the traditional rules of voice-leading. We will use three of these principles for motivating simple melodic properties.

The *Toneness Principle* uses the fact that pitch perception is more accurate in a middle pitch region centred approximately around D4. Therefore, it is assumed that melodies are more easily perceived in this region. The region also coincides with the mean pitch of children's and women's voices (and close also to the typical preferred register for male pop singers). It can be argued that this principle reflects the conceptual connection between singing and melody mentioned above. It is also in agreement with the mean pitch in a large sample of melodies that was found to be D#4 (see Huron, 2001). In Figure 4 the mean pitch for the main melody and for the remaining tracks are shown. It is clear that the main melody roughly follows the toneness principle. However,

the position of the peak is approximately C5, thus, considerably higher than the toneness centre possibly reflecting the simple principle of the melody being the top voice. The histogram of the main melody shows an approximate Gaussian shape, indicated by the fitted curve. The rest of the tracks are considerably more spread out, in particular towards the bass, but there might also be a few more cases of higher mean pitch indicating that it is not always the main melody that is the top voice.

The *Principle of Temporal Continuity* states that the formation of an auditory stream is stronger when the sound is continuous rather than consisting of brief sounds surrounded by rests. According to the principle, the longest possible pause between two events in one stream is roughly 800 ms. This is motivated from the research of echoic memory and stream segregation (van Noorden, 1975; Bregman, 1990). This principle implies that sustained timbres (voice, winds, strings) are preferred over percussive instruments (piano, marimba) and that legato is preferred over staccato, which also is in accordance with the aforementioned voice-melody connection. There are often rather long pauses between motives in a melody within the music examples in the database. Since these pauses mark structural boundaries in the melody, they should not be considered as critical with respect to the limit of 800 ms. On the contrary, the 800 ms limit might be used for marking the structural boundaries of the melody. One might assume that the accompaniment serves more as a rhythmic foundation rather than bringing out melodic features. The accompaniment would then not need to use the temporal continuity principle suggesting a difference between melody and accompaniment. This seems to be confirmed both regarding timbre and articulation. Sustained harmonic timbres are used for 84% of the main melodies while the same number for all other tracks amounts to only 43%. The mean articulation for the main melodies and for the remaining tracks is shown in Figure 5. There is a tendency for the main melody to be more legato articulated as indicated by the fewer cases of articulation

lower than 0.5 in the left graph. However, as indicated by the highest bar in both graphs there is a large amount of tracks that is performed legato in both cases, thus reducing the predictive power.

The *Pitch Proximity Principle* suggests that in order to maintain one melodic stream the successive intervals must be relatively small, reflecting the gestalt principle of good continuation. This is supported by the research in stream segregation (e.g. van Noorden, 1975; Bregman, 1990) as well as coinciding with statistical measurements in large melody collections. The accompaniment, if serving a more rhythmic or harmonic function, will have less obligations to be perceived as continuous streams. This is confirmed in the current data, as shown in Figure 6, which plots the histogram of the average pitch interval regardless of direction for each track. The intervals of the main melody lie in a rather narrow range with a peak around 2 semitones. This may very well also be regarded as reflecting the conceptual linkage between melody and singing, regarded as an optimization for vocal performance. The other tracks in the scores are generally considerably more heterogeneous with regards to successive intervals, including both larger and smaller

averages. Note that there are no main melodies in our sample that contain only one pitch which is quite natural from the point of view of melody as a narrative; pitch change makes it easier to identify/represent longer structures than just rhythm, hence supporting the story-telling.

We also suggest a principle of *Temporal Sensitivity* that corresponds to the toneness principle above assuming that melodies are more easily perceived when tone durations are within a region centred around 250 ms. This can be motivated by the measurements of just noticeable difference (JND) of duration which exhibit the highest sensitivity in this range (Friberg & Sundberg, 1995). In Figure 7, the JND for one note in an isochronous sequence is plotted as a function of IOI. It is not obvious whether one should consider JND with respect to the absolute deviation in milliseconds or the relative deviation in percent. Musical structures generally rely on categorical perception of relative duration rather than absolute duration, which is generally related to tempo cognition. Thus, the relative deviation should be the most relevant measure (e.g. Desain & Honing, 2003). However, as indicated in Figure 7, the relative JND

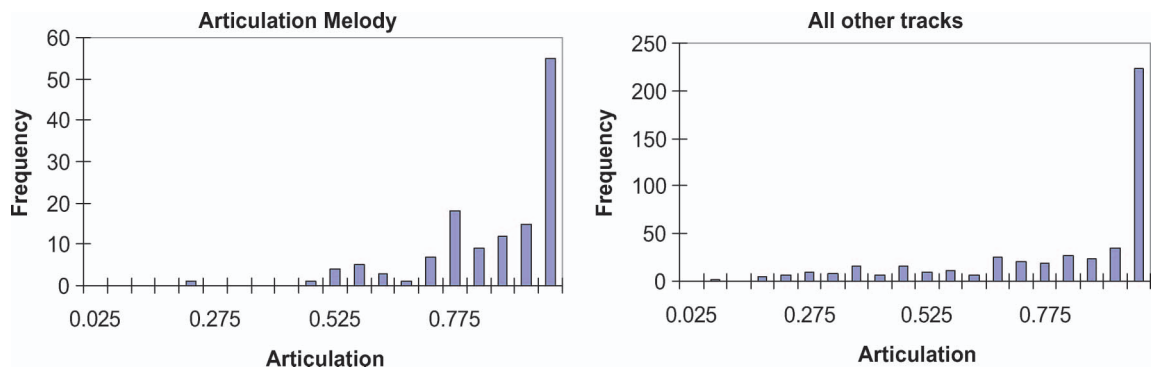


Fig. 5. Histogram of average legato/staccato for the main melody (left) and for all other tracks (right) in the training set. The articulation parameter is defined as duration/IOI thus legato means that the articulation parameter is equal to one.

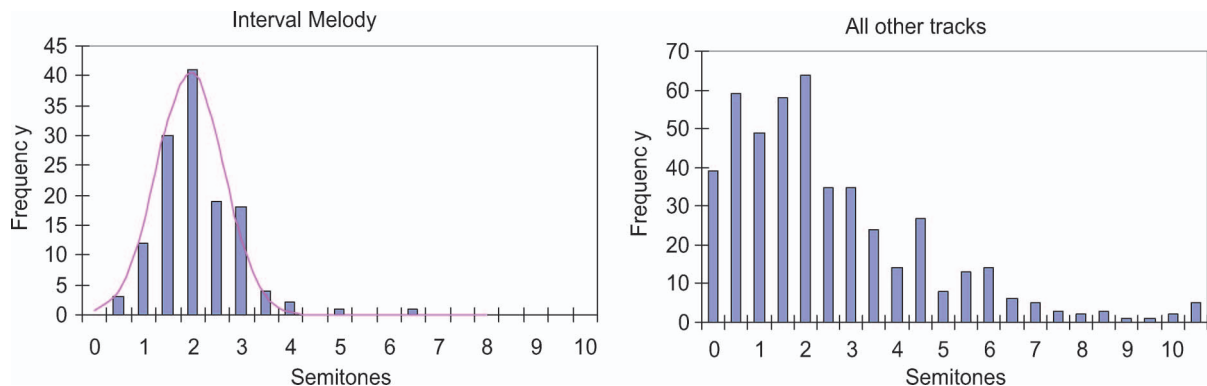


Fig. 6. Histogram of average pitch interval size in semitones for the main melody (left) and for all other tracks (right) in the training set.

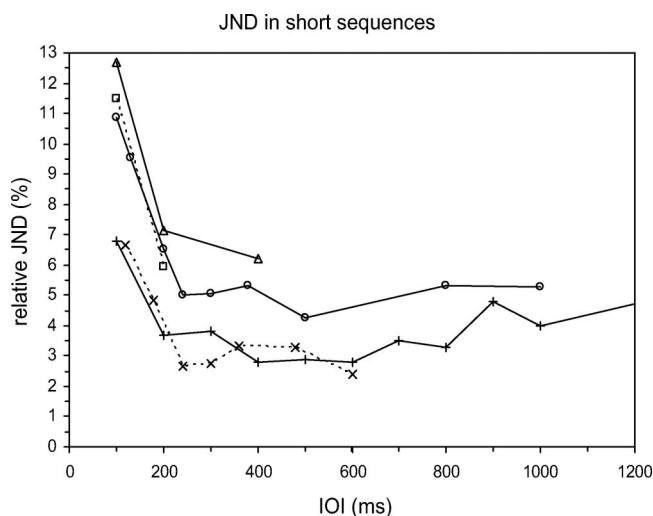


Fig. 7. JND in short isochronous sequences collected from five different studies (from Friberg & Sundberg, 1995).

increases suddenly for IOIs shorter than about 250 ms. In this lower region, the JND is constant in absolute terms and amounts to approximately 10 ms. This may indicate that the relativity of music timing is less effective for shorter durations (see also Friberg & Sundström, 2002; Desain & Honing, 1994). It also indicates that there are two different processes responsible for timing measurements. If both processes are considered, the combined JND has its minimum at the knee in the curve, that is, around 250 ms.

Statistical data for the IOI distribution of notes are rare in the literature, possibly due to the lack of databases of performed music. However, Huron (2001) with a sample of 13,178 notes but for a rather limited stylistic range, obtained a peak around 250 ms. In addition we investigated the IOI distribution for the current database using all examples. The relative frequency of IOIs is shown in Figure 8. Despite a large number of notes included (16,782) there were relatively few unique values. Therefore, the histogram was produced by weighting each IOI with a hanning window centred at each bin position and with a size 4 times larger than the bin step size. The solid line shows the resulting frequency with the bin step size of 50 ms, while the dotted line represents a 25 ms step size (the latter compensated by a factor of 2 in the plot). With the larger step size a clear single peak appears at 250 ms. However, the smaller step size also indicates a secondary peak at about 150 ms. Note that there is also a sharp break at 100 ms supporting previous evidence that this is the shortest note to be perceived on its own (Friberg & Sundström, 2002; London, 2004).

The typical IOI value of 250 ms is easily motivated from statistical investigations of tempo if we assume that the most important note value for a melody is half the

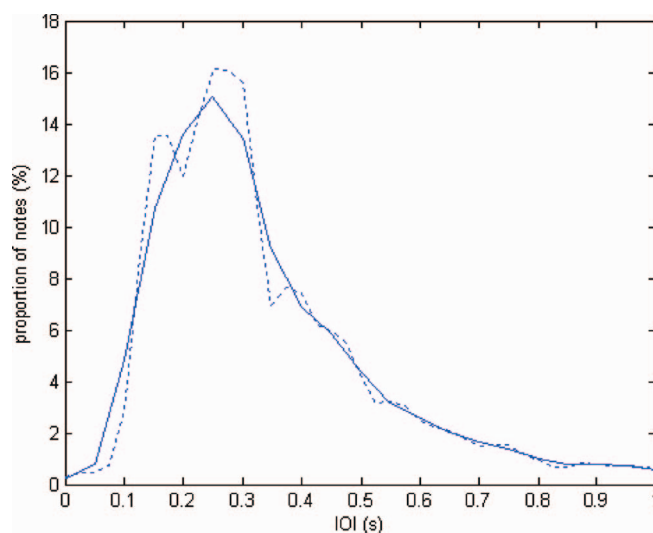


Fig. 8. Relative proportion of IOI for all notes in all melodies in the database.

tactus duration. Van Noorden and Moelants (1999) obtained a tempo distribution very close to the curve shown in Figure 8 but with the peak at 500 ms (see also Moelants, 2002). Many other studies suggest that 250 ms marks a boundary between different kinds of perceptual processing, see London (2004) for an overview.

For the purpose of applying the properties of IOI for discriminating between main melody and accompaniment we need to use average values over the tracks. Note that the skewness of the distribution in Figure 8 makes the mean value of the IOIs higher than 250 ms. Also the distribution of mean values is skewed in the same way. The skewness is reduced if the mean values are inverted, thus represented as notes per second instead. In Figure 9, the distribution of notes per second (inverted mean IOI) for the main melodies and for the remaining tracks is shown for the training data set. It is clear that the main melody is centred around a rather narrow range with a Gaussian-like distribution while the remaining tracks are considerably more spread out.

In addition to the features mentioned above, we hypothesized that a few additional low-level features may serve as candidates for predicting main melody. They were (1) normalized pitch mean across each score. This would be effective if the main melody is the top voice. This is often referred to as the skyline model and has been used for melodic matching in polyphonic music data bases (e.g. Uitdenbogerd & Zobel, 1998; Wiering et al., 2009). However, it would not be very effective as a single selection measure in the current investigation since only 50.4% of the music examples in the training set have the main melody on top. (2) Average sound level. Obviously the loudest voice will be perceptually salient and thus a potential candidate. (3) The total sounding duration of a voice normalized over each music example

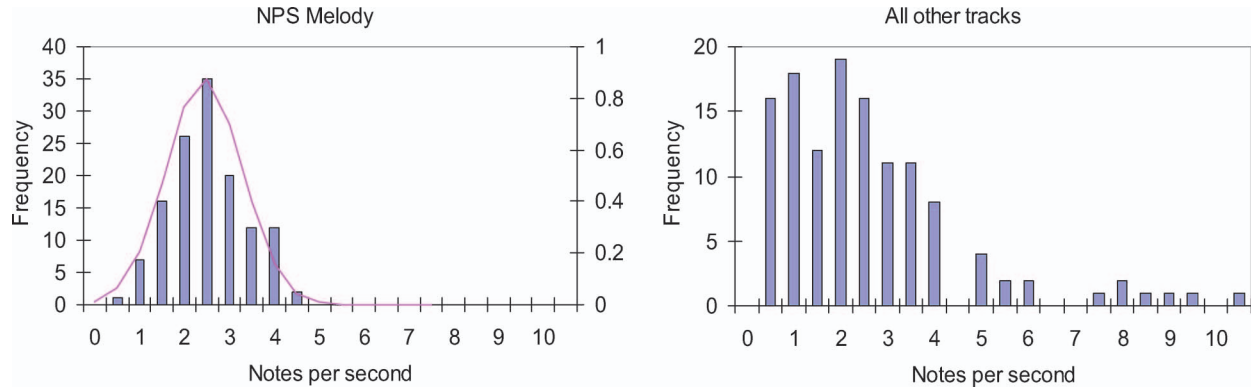


Fig. 9. Histogram of average notes per second for the main melody (left) and for all other tracks (right) in the training set.

was added assuming that the melody would have a substantial duration but lower than the total duration of the score. Accompaniment tracks such as the bass line can be assumed to have a longer duration, while some occasional accents would be shorter. (4) The relative proportion of simultaneous tones in each voice. The idea was that tracks consisting mainly of chords are more likely part of the accompaniment.

3.2 Measure of narrativity

The underlying assumption behind the narrativity measure is that a melody can be regarded as analogous to a story, in the sense of an intelligible whole comprised of a sequence of phrases with unique content. As mentioned above, the basis for this analogy is that the melody concept in a western cultural context seems to be closely linked to the concept of song, which can be regarded as a story told with both words and music. (In a historical as well as a global perspective, singing a story and telling a story are not always separate concepts; the same word often designates both activities.) Thus, in analogy with language, like a story is not just a string of words or a rigmarole, a melody can be regarded not just as a string of notes, but rather as a sequence of phrases that connect to each other to form an intelligible whole.

This reasoning implies a gestalt analogy between the concept of story and the concept of melody; the more evident gestalt qualities a story or melody exhibits, the easier it should be to grasp. According to the *prägnanz* gestalt principle (see e.g. Wertheimer, 1923/1938) a gestalt such as a story or melody should ideally have as simple structure as possible (which implies as few substructures as possible) and at the same time be as recognizable and easy to distinguish as possible, which implies a contrasting but coherent inner structure.

We have applied this concept to obtain a measure of how ‘melodic’ in the sense of ‘narrative’ a certain midi track is. The principle of the measure is simple: the greater the proportion of unique and pertinent melodic phrases in relation to the whole length of the track the

more ‘narrative’ or ‘melodic’ the track is expected to be perceived. This is obtained by applying the phrase analysis system developed by Ahlbäck (2004, 2007) on the individual tracks and collecting the phrases that have (1) the longest unique content, i.e. phrases whose series of events are not contained in another phrase to a higher degree and (2) that are not classified as mere repetitions of previous phrases, i.e. that constitutes a structurally significant difference to previous phrases (Ahlbäck 2007).

The measure is the proportion of the duration of unique phrases in relation to the total duration of the track (*phrase_ratio*) weighted against the proportional note content of the track (*note_ratio*), according to the formula

$$\text{Narrativity} = \frac{2}{3} \cdot \text{phrase_ratio} + \frac{1}{3} \cdot \text{note_ratio}. \quad (1)$$

This gives a number between 0 and 1 corresponding to the amount of narrativity. This means that the influence of the *phrase_ratio* is related to how much the track actually sounds. The weights reflect the gestalt principle of differentiation, which implies that the simplest level of gestalt organization is two contrasting substructures of equal prominence.

For a typical accompaniment part where the entire structure is composed by short repeated patterns with negligible differences the *phrase_ratio* becomes quite low, while in a more typical melodic structure with contrasting structures it becomes higher.

Figure 10 shows the output phrase analysis of a midi track with the unique phrase (A1) enclosed by a rectangle. The differences between the phrases marked A1 on the one-bar level are regarded negligible according to the phrase analysis, which gives a *phrase_ratio* of 1/8, which means that the unique phrase covers 1/8 of the total duration of the track.

Figure 11 shows the output phrase analysis of a more melodic midi track, where the unique phrases (two-bar level) are enclosed by rectangles. The differences in the repetition A1–A2 (on two-bar level) is considered



Fig. 10. Output result from phrase analysis of a midi track from 'English summer rain' by Placebo. Unique phrase shown by enclosed rectangle.

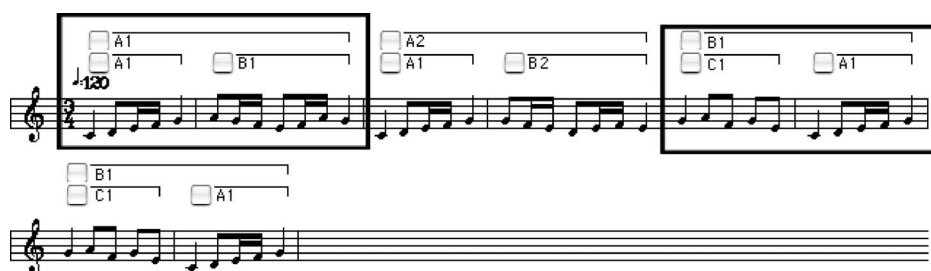


Fig. 11. Output result from phrase analysis of a midi track from a Swedish folk melody. Unique phrases shown by enclosed rectangle.

negligible by the analysis. In this case the unique phrases cover half the duration of the track giving a *phrase_ratio* of 1/2. Note that even though the same sub-phrase (A1 one-bar level) appear in both phrases the differences are considered significant. For details regarding the similarity classification see Ahlbäck (2007).

The narrativity measure in this sense is not in itself enough for measuring melodicty, since the other features discussed in this article, such as registral continuity expressed by the pitch proximity principle mentioned above, also contributes to the structural coherence, hence narrativity, of the melody.

4. Method

4.1 Music database

A set of music examples consisting of 242 polyphonic scores were collected from various commercial databases of ringtones for mobile phones. They were all encoded in the MIDI format. The majority of the music examples were instrumental versions of popular songs. The MIDI format is rather flexible in that it allows the music both to be encoded similar to the score with quantized note durations positioned in a metric grid or to be encoded just as individual note durations as a music performance. The music producer is thus free to just insert the notes at any time position, for example by playing the different tracks on a keyboard. Given the diverse origin of these examples, no general information about the score could

be assumed and the music could only be treated as a sequence of notes divided into different MIDI tracks.

All tracks were independently annotated by both authors into the categories main melody, melody, accompaniment, bass, drums, and effect sounds. The same category could be used for more than one track including the main melody category. The *main melody* category was used for the main melodic theme of the piece, likely to be the original sung melody. Doublings and octave transpositions were also included. The *melody* category was used for other tracks with a predominately melodic character. The *accompaniment* category was used for tracks with a predominantly harmonic (including chords) character or consisting of small rhythmic repeated figures.

The classification of the drums was trivial and was excluded in the subsequent analysis. Among a total of 1159 tracks 81% were assigned to the same category by both annotators. This is an indication that the given task was not as simple as expected. Within the group that was categorized the same, the distribution was the following: main melody 27%, melody 10%, accompaniment 39%, bass 23%, and effect sounds 1%. Among the group of tracks assigned to different categories, a common difference in interpretation was between main melody and melody. However, most combinations of categories did occur, including bass/effect sounds and main melody/bass.

The main annotated measure was a dichotomously coded variable that was 1 if a track was annotated as main melody by both annotators and 0 for all other

cases. This was used as the dependent variable in the subsequent analysis and denoted *main_mel*.

In addition, a measure of *annotated melodic weight* was estimated by weighting the annotated categories according to their assumed melodicity with the following weights: main melody 1, melody 2/3, accompaniment and bass 1/3, effect sounds 0. The average weight for the two annotators was the final measure. This measure was then less sensitive to small differences in the classification and was assumed to potentially improve the performance of a continuous prediction model.

The database was further randomly divided in two sets of equal size; a training set and an evaluation set, each containing 121 examples. All the subsequent analysis and optimization were done on the training set while the evaluation set was kept until the final cross-validation of the model.

4.2 Features

The features used in the experiment are listed in Table 1. The Pitch feature is directly reflecting the toneness principle and the Pitch interval is reflecting the pitch proximity principle. Normalized pitch over score was added reflecting the ‘melody on the top’ principle with the value 1 denoting the top voice. The Pitch interval variation was added assuming that a melody needs some variation to be narrative and distinguishable.

The IOI and IOI variation features reflect the temporal sensitivity principle and the Articulation and Timbre features reflect the principle of temporal continuity.

The Timbre measure was estimated from the general MIDI program specification and coded as 1 for sustained

harmonic instruments, 0.5 for percussive harmonic instruments, and 0 for non-harmonic sounds.

Sound level normalized over score is reflecting ‘the loudest—the most salient’ principle. No direct measurement of sound level was available since dynamics in a MIDI file is arbitrarily coded in terms of MIDI velocity. Therefore, a translation from velocity to sound level was done using measurements from a typical MIDI synthesizer (Bresin et al., 2002). Unfortunately, many of the examples did not have any coded variations in MIDI velocity implying that the Sound level feature did not contribute anything for these cases.

The Total sounding duration of each track was computed by summarizing all note durations and then normalized over each score.

The polyphony feature was a measure of the relative amount of simultaneous notes (chords) in each track with the value 1 for a track that only contains chords and the value 0 for a purely monophonic track.

The computed narrative feature was also normalized over score. This normalization over score for several features further emphasized the relative contextual effect of each feature as discussed above.

The features using at least two onsets (Pitch interval, Pitch interval variation, Articulation) were computed over continuous parts of the melody only. If a pause between notes was longer than 800 ms it was considered as a break in the melody and the features were not computed across this position.

4.3 Feature transformation

As shown in the previous section, many of the features appear in a rather narrow region closely resembling a

Table 1. List of features used for the prediction of the main melody.

Feature	Measure	Predicted optimal range	Gaussian Transf.	Symbol
Pitch	mean pitch in semitones (MIDI number)	Approx F2-G5	X	f0_mean_tr
Pitch normalized	Pitch mean normalized over score (range 0–1)	High		f0_mean_norm
Pitch interval	mean absolute interval to next note (semitones)	Small but > 0	X	df0_mean_tr
Pitch interval variation	standard deviation of pitch interval	Intermediate	X	df0_sd_tr
IOI	mean IOI (ms)	Intermediate	X	ioi_mean_tr
IOI variation	standard deviation of IOI	Small but > 0	X	ioi_sd_tr
Articulation	Legato/staccato quantification (range 0–1 with 1 for legato)	Legato		art_mean
Timbre	Sustained harmonic (1), percussive harmonic (0.5), nonharmonic (0)	Sustained harmonic		timbre
Sound level	Mean sound level normalized over score (0–1)	High		sl_mean_norm
Total duration	Total sounding duration normalized over score (range 0–1)	Intermediate		dr_sum_norm
Polyphony	Amount of simultaneous notes normalized over score (0–1)	Low		chord_percent
Narrative	The narrative measure normalized over score (0–1)	High		narrative_norm

Gaussian shape when plotted for the main melodies. For the rest of the tracks, these features appear both above and below the region defined by the main melodies. This implies that these features will have small predictive power when they are used in a linear prediction method, such as multiple regression. In fact, a feature can be evaluated as insignificant while its ‘true’ impact is rather high. Each feature F which was expected to have a certain region to be important for melody was therefore transformed to F_{TR} using a Gaussian curve that was fitted to the main melodies of the training set using the formula

$$F_{TR} = \exp\left(\frac{-(F - m)^2}{2\sigma}\right), \quad (2)$$

where m and σ are the fitted mean and standard deviation. The transformed data is then limited to the range 0 to 1. A high value of F_{TR} means that the feature is close to the centre of the distribution of the main melody while lower values means that it is either above or below the centre. This was applied to the features Pitch, Pitch interval, Pitch interval variation, IOI, and IOI variation, as marked in Table 1 and indicated by the ending ‘_tr’ in the feature symbol.

4.4 Data pre-selection

The lowest track, as defined by the average pitch feature, was found to never represent the main melody. It was in most cases part of the bass line and was therefore automatically excluded both from the training and the evaluation.

In addition, a few outliers were identified in the group of the main melody in the training set. There were some doublings of the main melody with very low sound level which removed the effect of the sound level feature. These tracks were manually excluded during the training of the model but not during the final evaluation.

4.5 Prediction method

The primary method that was applied for predicting the main melody was multiple regression. The reason for this choice was that it is a straightforward method which has been thoroughly investigated and used in similar experiments. It can also be applied both to categorical and continuous data, thus the two different dependent measures above can directly be compared. The maximum of the regression equation was taken as the predicted main melody for each score. The final hit rate was computed by counting how many music examples in which the main melody was predicted correctly.

Looking at the problem as a binary classification task (main melody/not main melody) a multitude of alter-

native methods is available. A natural extension is logistic regression which allows an easy comparison with the standard multiple regression. In addition, the more contemporary classification method Support Vector Machines (SVM) was also applied. SVM basically extends the linear behaviour of multiple regression by transforming the features to a higher dimensional space and then finding a linear solution (rather a hyperplane) which is then transformed back to the original space. This means that the operatic regions for the features above can potentially be found even without the Gaussian transformation. It also considers feature interactions. Thus, potentially, SVM should be able to find a better fit given that the features and dependent functions are well defined.

5. Results

The cross-correlation between all the features and the dependent variable, i.e. the annotated main melody is shown in Table 2. As indicated in column 1 all features except Articulation have a reasonable correlation with the main melody. All of the correlations are also in the right direction, including *chord_percent* that is expected to be negative for the main melody. Thus, most of the features have a potential prediction power. Although many of the cross-correlations are significant, most of the correlations are fairly low. The highest correlations are shown in bold (above 0.50) and are usually different measures of the same basic features, such as between Interval and Interval variation.

For comparison, Table 3 shows the correlation of the untransformed features with *main_mel*. If we compare these values with the corresponding transformed features in Table 2 we observe that the transformation has a large effect on the size of the correlation in most cases. In three cases it goes from non-significant to highly significant (*df0_mean*, *df0_sd*, *ioi_sd*), for *ioi_mean* it improves somewhat, while for the *f0_mean* it goes in the wrong direction. The latter could be reflecting that the principle of the top voice is nevertheless valid. This is also indicated in the higher correlation of *f0_mean_norm* compared to *f0_mean_tr* in Table 2.

5.1 Multiple regression

A solution for multiple regression was found for the training set using the data described above. From the obtained regression equation coefficients the main melody was predicted by selecting the track with the maximum predicted value for each score. Thus, only one main melody was predicted for each score. If the predicted main melody was also marked as the main melody the prediction was successful for that score. The prediction of the evaluation set was done using the

Table 2. Cross-correlation of features and the dependent variable *main_mel*. All correlations above 0.50 are shown in bold. Significance levels are marked with * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

	<i>main_mel</i>	<i>f0_mean_tr</i>	<i>f0_mean_norm</i>	<i>df0_mean_tr</i>	<i>df0_mean_norm</i>	<i>df0_sd_tr</i>	<i>df0_sd_norm</i>	<i>ioi_mean_tr</i>	<i>ioi_mean_norm</i>	<i>ioi_sd_tr</i>	<i>art_mean</i>	<i>dr_sum_norm</i>	<i>timbre</i>	<i>sl_mean_norm</i>	<i>chord_percent</i>
<i>f0_mean_tr</i>	0.26***														
<i>f0_mean_norm</i>	0.35***	0.58***													
<i>df0_mean_tr</i>	0.34***	0.20***	0.18***												
<i>df0_sd_tr</i>	0.25***	0.15**	0.09	0.50***											
<i>ioi_mean_tr</i>	0.27***	0.02	0.04	0.05	0.03										
<i>ioi_sd_tr</i>	0.28***	0.08	0.03	0.24***	0.15**	0.56***									
<i>art_mean</i>	0.06	0.00	-0.02	0.19***	0.21***	-0.32***	0.02								
<i>dr_sum_norm</i>	0.14**	-0.07	-0.09	0.09	0.12*	-0.24***	0.00	0.60***							
<i>timbre</i>	0.27***	0.06	0.13**	0.34***	0.19***	-0.00	0.07	0.13**	0.09						
<i>sl_mean_norm</i>	0.47***	0.04	0.07	0.12*	0.08	0.15**	0.16**	-0.08	-0.02	0.01					
<i>chord_percent</i>	-0.34***	-0.04	-0.03	-0.12*	-0.14**	-0.15**	-0.13**	-0.12*	0.03	-0.10*	-0.30***				
<i>narrative_norm</i>	0.43***	0.16***	0.15**	0.34***	0.24***	0.05	0.25***	0.18***	0.28***	0.20***	0.26***	-0.23***			

coefficients from the training set both for the multiple regression and for the feature transformation. The overall results are shown in Table 4. In the training set only 9 scores out of 121 could not be correctly predicted and slightly more, 11 scores, for the evaluation set. Thus, the results appear to be quite stable for unknown data. Further optimization of the hit rates could be problematic since there are very few cases which the prediction misses. That would require a larger data base.

The details of the regression data for the training set are shown in Table 5. Of the 12 features, eight gave a significant contribution to the regression. Some of the hypothesized features, such as *f0_mean_tr* did not contribute to the final prediction. In this case it appears that the principle of the top voice as reflected by the feature *f0_mean_norm* was a stronger predictor than *f0_mean_tr*. The squared semipartial correlations sr^2

Table 3. Correlation of the un-transformed features with the dependent variable *main_mel*. Significance levels as in Table 2.

	Main mel
<i>f0_mean</i>	0.28***
<i>df0_mean</i>	-0.08
<i>df0_sd</i>	0.02
<i>ioi_mean</i>	-0.19***
<i>ioi_sd</i>	0.00

Table 4. The overall hit rate of predicting the main melody for the optimization and evaluation set.

	Total number of scores	Hits	Percentage
training set	121	112	92.6
evaluation set	121	110	90.9

Table 5. Multiple regression coefficients for each feature. Significance levels as in Table 2.

feature	β	sr^2	p -value
<i>f0_mean_tr</i>	0.058	0.046	0.176
<i>f0_mean_norm</i>	0.235	0.188	0.000***
<i>df0_mean_tr</i>	0.087	0.069	0.045*
<i>df0_sd_tr</i>	0.046	0.039	0.255
<i>ioi_mean_tr</i>	0.207	0.155	0.000***
<i>ioi_sd_tr</i>	0.012	0.009	0.788
<i>art_mean</i>	-0.056	0.040	0.240
<i>dr_sum_norm</i>	0.193	0.145	0.000***
<i>timbre</i>	0.145	0.134	0.000***
<i>sl_mean_norm</i>	0.319	0.292	0.000***
<i>chord_percent</i>	-0.145	0.130	0.000***
<i>narrative_norm</i>	0.146	0.126	0.000***

represent the individual contribution from each feature when interdependence effects are removed. We see that the simple and intuitive measures *f0_mean_norm* and *sl_mean_norm* give the strongest contribution. The *sr*² values are only slightly smaller than the beta coefficients indicating that the amount of potentially problematic interdependence among the features is small. Note that all the features that are normalized over score are highly significant, supporting the relative, context-dependent notion of melody.

The individual contribution of each feature was also evaluated in terms of the decrease in hit rate counted over scores when one feature was removed and the model was computed for the remaining features. This is similar to the squared semipartial correlations but computed for the scores instead of individual tracks. Due to the few number of cases, it was necessary to compute this for the whole data set. The resulting decrease in hit rate is shown in Table 6. The effect when removing one feature is small and the maximum decrease in hit rate is about 3%. Thus, the remaining features are able to compensate for the missing information to some extent. For example, when *f0_mean_norm* is removed, *f0_mean_tr* increases in importance and becomes highly significant instead. The changes in hit rate follow approximately the *sr*² coefficients in Table 5. One exception is the narrative measure that, when removed, generates a quite small decrease in hit rate. However, due to the small number of cases these results should be treated with some caution.

In addition, both robust regression and logistic regression were tested using the same procedure. The results were very similar with exactly the same hit rate for the training set for all three methods and with similar *p*-values.

Table 6. Decrease in hit rate for the whole data set when one feature is removed and the model is computed on the remaining features. The cases are presented relative the total number of hits for all features (224 correct cases out of 242, 92.6%), see Table 3.

Feature	Cases	% Change
f0_mean_tr	0	0
f0_mean_norm	−4	−1.7
df0_mean_tr	−2	−0.9
df0_sd_tr	0	0
ioi_mean_tr	−4	−1.7
ioi_sd_tr	0	0
art_mean	0	0
dr_sum_norm	−6	−2.5
timbre	−4	−1.7
sl_mean_norm	−8	−3.3
chord_percent	−1	−0.5
Narrative_norm	−1	−0.5

A few alternative features were compared within the multiple regression method. For example, the annotated melodic weight measure, described above, was also tried as the main dependent variable. However, this resulted in a slightly lower hit rate (90.1% for the training set), and was hence not used in the analysis. The use of notes per second for the tempo features was compared with the IOI measure and the latter was found to be slightly better.

5.2 Support vector machine modelling

The LIBSVM library was used for the SVM classification (Chang & Lin, 2001). The training and fine-tuning of the SVM model followed the guidelines given by Hsu et al. (2008). All the features were scaled to the range −1 to 1 and the RBF kernel function $K(x,y)=exp(-\gamma||x-y||^2)$ was used. The model parameters γ and *C* were optimized with a grid-search using cross-validation. That is, the parameters were systematically varied over a large range to find an area with good performance which was then scanned with a finer resolution. The same data set as used in the multiple regression was used for training.

The best solution gave an accuracy of 89.8% right classification of the included tracks (536 tracks out of a total of 597) for the training set. It was translated into hit rate counted over scores by using the probability output of the model in the same way as for the multiple regression. Remarkably, the resulting hit rate was 92.6% for the training set and 90.9% for the evaluation set, thus, exactly the same hit rates as obtained using multiple regression, see Table 4. It seems that the SVM solution converged to a simple linear solution without much higher order dependencies and interaction among the features. Theoretically, the untransformed features might have been a better starting point since the SVM method with this kernel will consider Gaussian distributions in higher dimensional space, thus potentially finding a better fit including interactions. However, surprisingly, the same feature set with untransformed features gave a slightly poorer fit for the SVM (90.1% for the training set), thus confirming a lack of interaction effects between the features.

5.3 Analysis of individual cases

It turned out that the initial examples in Figures 1–3 (that by chance were all grouped into the training set) were all correctly classified.

The wrongly predicted cases were further inspected in order to try to reveal the reason for prediction failure. For the training set, the wrongly predicted scores were exactly the same for the multiple regression and the SVM method. For the evaluation set only two scores were different. This indicates that the solution found for the SVM method was very similar to the simple linear regression solution. Most of the wrongly predicted tracks

were still rated high on the annotated melodicty measure meaning that it was at least annotated as ‘melody’ by both annotators. Thus, the errors could be considered minor in that it predicted an alternative melody rather than e.g. some accompaniment. There were a few annotation errors (such that one copy of the main melody was not properly labelled) but most of the problematic cases were scores in which the main melody shifted from one track to the other or scores with interwoven melodies in several tracks. Thus, cases which were problematic even for human annotators.

6. Conclusion and discussion

In summary, twelve different features were identified as potentially relevant for determining the main melody in a polyphonic score. They were related to pitch (four features), timing (three features), sound level, articulation, timbre, polyphony, and narrativity. Eleven of these significantly correlated with the annotated main melody. Of these, eight were still significant when combined in a multiple regression model. The simple intuitive features relative to pitch height in the score (*f0_mean_norm*) and sound level were found to explain a fair amount of the variation. Other strong features were the IOI range (*ioi_mean_tr*), polyphony (*chord_percent*), timbre, narrative, and the total sounding duration.

How can these results be interpreted? The relative pitch height measure and IOI range relate to the vocal analogy principle, which reflects that the melody is generally a vocal part in this database. Likewise, the polyphony measure reflects the concept of melody as a fundamentally monophonic phenomenon. The sound level reflects that the producers of the ring tones regard the main melody to be a significant element for the identification of a song, so does the timbre measure. The narrative measure and the total sounding duration reflect the narrative concept of melody (the latter in the sense that it reflects the degree to which the track is addressing the whole piece of music). Thus, it turns out that different features, related both to instrument/voice, production and musical structure concur in the indication of main melody in this sample.

The effect of the sound level feature was strong. Considering that this feature had some inconsistencies in the coding the effect was quite remarkable. It was not coded in all the examples and some doubled melody tracks had a very low sound level. Given a more consistent coding of the data base the effect of the sound level would possibly increase even more. The articulation on the other hand showed very little relation to the main melody in this data selection. One possibility is that articulation was not properly coded in the examples. For example, if a percussive sound is used, a short duration is obtained even if the MIDI note is long. For these

features it seems clear that the outcome of the prediction models may vary according to the particular selection of examples and how the features are coded.

The selection of features can always be discussed. For example, von Hippel (2000) suggested an interaction effect of tessitura and interval direction implying, for instance, that the probability for a downward motion is higher if it was preceded by an upward leap. However, the pitch features already now overlap; this is indicated by the exclusion of some pitch features in the final prediction. Therefore it is unlikely that such a refinement would yield any significant improvement of the overall hit rate.

The overall hit rate of about 90% is highly dependent on the complexity of the music examples in the used database. If the music examples were simple enough it would be easy to predict the main melody 100% correct. Therefore it would be interesting to apply the method to other databases, such as the one used by Rizo et al. (2006).

An interesting issue is to analyse the reasons for the wrongly predicted cases. The application of different methods with very similar results indicates that it is not due to the prediction method. The number of features seems also to be sufficient in that eight different features were enough and that exclusion of one feature (Table 5) made a small impact on the final prediction. Thus, the remaining features were able to compensate for the missing one. Still, it is possible that some essential aspects that can be formulated into a feature are missing. Another possible source of error is the annotation. The two annotators only fully agreed on 81% of all annotated categories, thus, lower than the hit rate of about 90%. This indicates that the problem is not completely determined in terms of ground truth.

The two different prediction methods Multiple Regression and SVM yielded very similar results in the end. In order to obtain the resulting hit rate both methods needed to be optimized by hand. For Multiple Regression it was important to transform the features and identify outliers. For SVM different methods had to be tested and the model parameters had to be optimized manually. This indicates that the rather simple method Multiple Regression can be quite adequate for this type of task given that the features with a restricted working range are modified using e.g. a Gaussian transformation.

The present features can be used for other purposes as well. An obvious extension is to use the features for judging melodic similarity, similar to the method used by e.g. Eerola et al. (2001). A trivial task would be to use the features for identifying multiple copies of similar tracks such as doublings of the melody in the current database.

An automatic recognition of the main melody in a MIDI file has some potential applications. The obtained hit rate of about 90% can be acceptable, in particular since the model often selects an alternative melodic track

when it misses the main melody. A previous version has been used in a commercial system for automatically enhancing ring tones with respect to musical quality and emotional expression.¹ The main melody in a score is enhanced (after being identified) by adding copies of itself with different timbres and transpositions according to the intended emotional expression (Bresin & Friberg, 2000). In addition, timing, sound level and articulation are changed using the KTH rule system for music performance (Friberg et al., 2006).

The main melody is often considered the most characteristic feature of a piece of music and the one that people remember. Melodic understanding would thus improve the matching of melodies in, for instance, query-by-humming systems as well as give a further understanding of the underlying musical properties that constitute the rather vague notion of melodiosity.

Acknowledgements

This work was done within the project SYMIR supported by the Swedish Research Council, Grant Nr. 2005–3696.

References

- Ahlbäck, S. (2004). *Melody beyond notes: A study of melody cognition* (PhD thesis). Göteborg University, Sweden (publications from the Department of Musicology).
- Ahlbäck, S. (2007). Melodic similarity as a determinant of melody structure. *Musica Scientiae, Discussion Forum*, 4A, 275–278.
- Bregman, A.S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.
- Bresin, R., & Friberg, A. (2000). Emotional coloring of computer controlled music performance. *Computer Music Journal*, 24(4), 44–63.
- Bresin, R., Friberg, A., & Sundberg, J. (2002). Director musices: The KTH performance rules system. In *Proceedings of SIGMUS-46*, Kyoto, pp. 43–48.
- Chang, C., & Lin, C. (2001). LIBSVM: A library for support vector machines. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Desain, P., & Honing, H. (1994). Does expressive timing in music performance scale proportionally with tempo? *Psychological Research*, 56, 285–292.
- Desain, P.W.M., & Honing, H.J. (2003). The formation of rhythmic categories and metric priming. *Perception*, 32(3), 341–366.
- Eerola, T., Järvinen, T., Louhivuori, J., & Toiviainen, P. (2001). Statistical features and perceived similarity of folk melodies. *Music Perception*, 18, 275–296.
- Friberg, A., & Ahlbäck, S. (2006). A method for recognising the melody in a symbolic polyphonic score. In *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC)*, Bologna, Italy, Abstract.
- Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance. *Advances in Cognitive Psychology, Special Issue on Music Performance*, 2(2–3), 145–161.
- Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America*, 98(5), 2524–2531.
- Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19(3), 333–349.
- Hsu, C.W., Chang, C.C., & Lin, C.J. (2008). A practical guide to support vector classification. Retrieved October 2, 2008, from <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>
- Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19(1), 1–64.
- Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.
- Kennedy, M. (1980). *The Concise Oxford Dictionary of Music*. London: Oxford University Press.
- London, J. (2004). *Hearing in Time: Psychological Aspects of Musical Meter*. New York: Oxford University Press.
- Madsen, S.T., & Widmer, G. (2007, January 6–12). A complexity-based approach to melody track identification in MIDI files. In *Proceedings of the International Workshop on Artificial Intelligence and Music (MUSIC-AI 2007)*, Hyderabad, India.
- Moelants, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition*, Sydney (pp. 580–583). Adelaide: Causal Productions.
- Perricone, J. (2000). *Melody in Songwriting: Tools and Techniques for Writing Hit Songs*. Boston: Berklee Press.
- Piston, W., & DeVoto, M. (1987). *Harmony* (5th ed.). New York: W.W. Norton.
- Rizo, D., Ponce de León, P.J., Pertusa, A., Pérez-Sachno, C., & Iñesta, J.M. (2006). A pattern recognition approach for melody track selection in MIDI files. In *Proceedings of the 7th International Symposium on Music Information Retrieval ISMIR 2006*, Victoria, Canada, pp. 61–66.
- Uitdenbogerd, A.L., & Zobel, J. (1998). Manipulation of music for melody matching. In *Proceedings of the ACM Multimedia Conference*, Bristol, UK, pp. 235–240.
- van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, 28(1), 43–66.
- van Noorden, L.P.A.S. (1975). *Temporal Coherence in the Perception of Tone Sequences*. Eindhoven: Institute for Perception Research.
- von Ehrenfels, C. (1890). Über Gestaltqualitäten. *Vierteljahrsschrift Für Wissenschaftliche Philosophie*, 14, 249–292.

¹see www.notesenses.com

- von Hippel, P. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception*, 17(3), 315–327.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung*, 4, 301–350. (Excerpts translated into English as ‘Laws of organization in perceptual forms’, in *A Source Book of Gestalt Psychology*, pp. 71–88, by W.D. Ellis, Ed., 1938, New York: Hartcourt, Brace and Co.)
- Wiering, F., de Nooijer, J., Volk, A., & Tabachneck-Schijf, H.J.M. (2009). Cognition-based segmentation for music information retrieval systems. *Journal of New Music Research*, 38(2), 139–154.