Int. J. Appl. Math. Comput. Sci., 2020, Vol. 30, No. 4, 761–772
DOI: 10.34768/amcs-2020-0056

AI BASED ALGORITHMS FOR THE DETECTION OF (IR)REGULARITY IN MUSICAL STRUCTURE

LORENA MIHELAČ [a], JANEZ POVH [b,c,*]

[a] SciDrom Scientific Lab
School Center Novo Mesto
Šegova ulica 112, 8000 Novo Mesto, Slovenia
e-mail: lorena.mihelac@sc-nm.si

[b] Faculty of Mechanical Engineering
University of Ljubljana
Aškerčeva ulica 6, 1000 Ljubljana, Slovenia
e-mail: janez.povh@fs.uni-lj.si

[c] Institute of Mathematics, Physics and Mechanics
Jadranska ulica 19, 1000 Ljubljana, Slovenia

Regularity in musical structure is experienced as a strongly structured texture with repeated and periodic patterns, with the musical ideas presented in an appreciable shape to the human mind. We recently showed that manipulation of musical content (i.e., deviation of musical structure) affects the perception of music. These deviations were detected by musical experts, and the musical pieces containing them were labelled as irregular. In this study, we replace the human expert involved in detection of (ir)regularity with artificial intelligence algorithms. We evaluated eight variables measuring entropy and information content, which can be analysed for each musical piece using the computational model called Information Dynamics of Music and different viewpoints. The algorithm was tested using 160 musical excerpts. A preliminary statistical analysis indicated that three of the eight variables were significant predictors of regularity (`E_cpitch`, `IC_cpintfref`, and `E_cpintfref`). Additionally, we observed linear separation between regular and irregular excerpts; therefore, we employed support vector machine and artificial neural network (ANN) algorithms with a linear kernel and a linear activation function, respectively, to predict regularity. The final algorithms were capable of predicting regularity with an accuracy ranging from 89% for the ANN algorithm using only the most significant predictor to 100% for the ANN algorithm using all eight prediction variables.

Keywords: regularity, musical structure, perception, AI algorithms.

1. Introduction

1.1. Motivation. Regularity exists in natural and human-made objects, including biology, physics, engineering, architecture, and art, and plays an important role in human life. The detection of repeated structures (patterns) is important, as it governs our recognition and understanding of the world (Pauly *et al.*, 2008). Therefore, finding patterns that are repeated and form a regular structure can help understand and analyze abnormalities in the structure due to some criteria (e.g., unexpected use of chords in the harmonic progression and its impact on the listener enjoyment). Formally, we can define regularity as a subset X_R of the set of all configurations X which have some structure that an observer tends to utilize or recognize (Feldman, 1997, p. 3).

In music, regularity is experienced as a strong structured texture with dominant periodic patterns and strong neighboring relationships (Manjunath *et al.*, 2000), where musical ideas are arranged in a shape appreciable by the human mind (Pole, 2014). Conversely, irregularity is experienced in a non-structured or weakly structured

*Corresponding author

musical piece, where the relationship between patterns can rarely be detected, and the enjoyment is affected due to the large mental space required for processing the musical content full of novelties (Kramer, 1988).

In our recent study (Mihelač and Povh, 2020), we examined the listener acceptability of music based on the complexity of harmony and proposed three objective measures for complexity: (i) *complexity of harmony* measured based on the presence of basic tonal functions and parallels in the harmonic flow, (ii) *unigram and bigram entropy* measuring the complexity in the harmonic progression, and (iii) *regularity* in terms of the order (regularity) and disorder (irregularity) found in the harmonic progression.

The latter was detected by a human expert (considering also comments of two additional musical experts) based on the presence of particular circumstances in musical excerpts (e.g., the occasional appearance of chords in harmony, non-detectable functions in harmony due to the use of figurated chords, emphasized rhythm or melody capturing the attention of the listener and placing the harmony in the background, etc.). Specifically, if at least one of 10 situations described previously by Mihelač and Povh (2020) appears within the harmonic flow, the musical example was labelled as irregular; otherwise, it was regular. A dataset containing 160 musical excerpts with all three complexity measures evaluated, is available at `http://kt.ijs.si/data/DATA_HARMCOMP.zip`.

This process of detecting irregularity has two drawbacks: (i) it demands high expertise of a listener/analyst and (ii) it is difficult to repeat the results (i.e., different experts can differ regarding some marginal examples in the labels that they assign). This motivated us to search for artificial intelligence (AI) algorithms capable of replacing human experts to explain and predict the (ir)regularity of musical excerpts. We built support vector machine (SVM) and artificial neural networks (ANN) classifiers that accept as input eight variables measuring entropy and information content. These variables were selected, because they can be evaluated for each musical excerpt automatically by using the computational model called Information Dynamics of Music (hereafter referred to as IDyOM) and various features, i.e., viewpoints (Pearce, 2005; 2018).

1.2. Main contributions. The main contributions of this paper are as follows:

- We supplement the dataset from our other work Mihelač and Povh (2020) with eight new variables (features): `IC_cpitch`, `IC_cpint`, `IC_cpintfref`, `IC_cpitch⊗dur`, `E_cpitch`, `E_cpint`, `E_cpintfref`, `E_cpitch⊗dur`. The first four variables describe different viewpoints of information content, whereas the last four variables describe different viewpoints of entropy. These additional features were computed by IDyOM. The supplemental dataset (HARMCOMP_2) with MIDI files, and sheet music is available to the research community at `http://kt.ijs.si/data/DATA_HARMCOMP.zip`.

- We analysed the relevance of the new features for irregularity using statistical analysis, including a Mann-Whitney-Wilcoxon test, which showed that the distributions of `E_cpitch`, `IC_cpintfref`, and `E_cpintfref` in the dataset displayed statistically significant differences regarding the subgroups of regular and irregular music excerpts, therefore suggesting them as good candidates for predicting regularity.

- We built SVM and ANN models to predict regularity based on the eight new features. Using 10-fold cross-validation, we showed an accuracy ranging from 89 % (achieved with an ANN using only the most significant predictor) to 100 % (achieved with an ANN using all eight prediction variables). These results suggest that AI-based algorithms can potentially replace expert-based detection of musical (ir)regularity.

All statistical results were obtained using the software `R` version 3.6.0.

1.3. Glossary. In this subsection we present several musical terms used in this study that might be unclear to readers unfamiliar with music. *Melody* (Fig. 1) is a horizontal appearance of notes representing a unique ordering of notes from a specific scale (Solomon, 2019, p. 24) and having an organized and recognizable shape. On its lowest level at the music surface (Jackendoff, 2009), it can be defined as a sequence of events (hereafter referred to as $e_1,\ e_2,\ldots$), in which each event e_1 is associated with *pitch* (highness or lowness; the frequency of a sound), *duration* (the length of a sound), *loudness* (intensity; loudness or softness of a sound), and *timbre* (quality or colour of a sound).

Harmony is a vertical presentation of tones (Fig. 1), which represents the structure of music with respect to

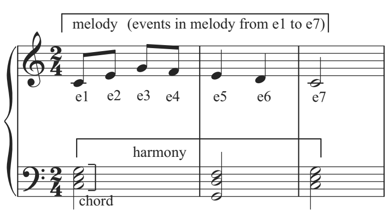

Fig. 1. Examples of melody and harmony.

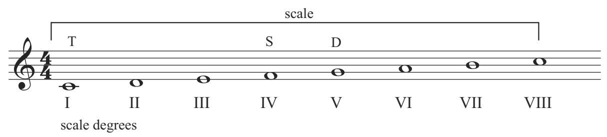

Fig. 2. Scale and scale degrees.

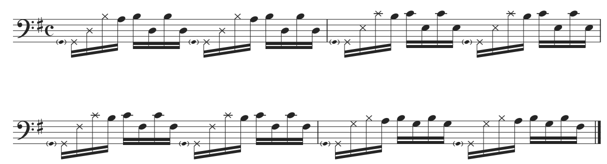

Fig. 3. Example of implied harmonies (Prelude in G major for cello solo, BWV 1007, from J.S. Bach).

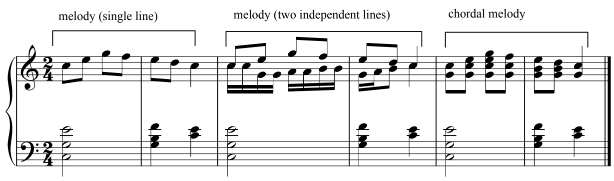

Fig. 4. Texture in melody: single line, two independent lines, and chordal melody.

chord composition and progression. *A chord* (Fig. 1) is a combination of at least three tones (sounds) performed simultaneously. A *non-chord tone* is a tone that does not belong to the prevalent (predominant) chord (Solomon, 2019).

Key designates a certain pitch as the tonal center of a musical piece and considers other pitches as scale degrees (the position of a note in a scale) around that center. *Scale* is a collection of discrete pitch relationships (Burns, 1999, p. 215) comprising a pattern of whole and half steps, whereas *scale-degree* indicates the position of a note within a scale (Solomon, 2019). The keys are simply named by the scale (see Fig. 2) on which they are based, e.g., (key) C major, C minor, etc. (Benward and Saker, 2008). *Function* in music does not have the same meaning as in mathematics. In music, a function describes the role a chord plays with respect to the root (tonic function) of a key. We can define in each scale (key) three *main* or *basic harmonic functions*: *Tonic* is the first scale degree (I) (i.e., the base of the key), *Subdominant* is the fourth scale degree (IV) of the key, and *Dominant* is the fifth scale degree (V) of the key.

All other scale degrees in a scale are defined as parallels (secondary degrees) of the three main functions (Dahlhaus, 2014): subdominant parallel (Sp; II degree), dominant parallel (Dp; III degree), and tonic parallel (Tp; VI degree).

Each tone in a melody can imply a certain harmony. In some melodies, the tones can be organized in such a way that merging them together results in a very strong *implied harmony* (Fig. 3). In this example, the noteheads, presented as crosses, form implied harmonies on the pedal point “G” (presented as noteheads between brackets), a tonic (T) in the first bar, an implied subdominant in the second bar, an implied dominant in the third bar, and again a tonic in the fourth bar.

Pedal point or *pedal tone* is a bass note (presented as notehead between brackets in Fig. 3), usually the tonic or dominant used through a sequence, including some chords (harmonies), which shifts around it.

In the continuation, we define the *texture* of the melody. In music, the texture of a melody is how the melodic material is combined in a composition and its effect on the overall quality of the sound in a musical piece (Benward and Saker, 2008). *Homophonic music* is music comprising melody and harmony. The texture of a melody can be presented in different ways (see Fig. 4). These can have either only one line or two or more lines and be written in a very independent way from the perspective of pitch and rhythm or as an example of chordal melody, when all the voices below the upper line have similar rhythmic material.

2. Scientific background and related work

During the listening process, the listener applies models acquired through the learning of regularities found in musical structure via exposure to music (Pearce, 2018), with this exposure either long-term (e.g., entire lifetime) or short-term (during the listening of a single composition).

During music processing, predictions of music are generated in order to organize and process the perceived musical content (Pearce, 2018). The ability to predict a subsequent event that meets listener expectations is important in the listening process, as it affects (among other things) the aesthetic experience. For example, if musical structure is somehow manipulated (e.g., by composer or performer), it creates in the listener the feeling of enjoyment if the event has happened or disappointment if not, tension if the event is delayed, or the feeling that the structure is ambiguous when the event is missing (Meyer, 1957; Narmour, 1990).

The idea that the structure of the musical content in a musical piece can affect the listener perception of a musical piece was proposed by Meyer in his seminal book *Emotion and Meaning in Music* (Meyer, 1957), where he outlined how some structures in musical pieces create higher or lower perceptual expectations for subsequent events depending on *how* the structure is manipulated by a composer.

The structure of a musical piece can be manipulated by deviating the form. Too complicated or even amorphous music, which includes a succession of heterogeneous ideas, without any relation between them affects the acceptance of music by a listener, as in the end, music should be structured in an understandable and (to a certain degree) predictable way in order to be properly enjoyed (Pole, 2014). The structure of musical dimensions, such as melody, harmony, and rhythm, can be deviated, as well. An unexpected chord in a harmonic progression, unexpected relationships between two chords in a chord progression that mismatches the rules of musical syntax, unexpected intervals in a melody, or unusual rhythm are examples of deviations (Rohrmeier, 2011; Rohrmeier and Pearce, 2018).

Listeners tend to use a set of basic perceptual principles that are applied to different musical styles and dependent on the type of music to which the listeners are exposed (Krumhansl, 2004). If the content of a musical piece does not meet these principles, it might confuse the listener based on the unusual use of musical dimensions or elements (e.g., unexpected use of pitch, intervals, rhythm, etc.) in the musical content. In such a case, the information provided is unrecognized (Finnas, 1989; Edmonds, 1995), and the *regularity*, sometimes posited as a mid-point between order and disorder (Grassberger, 2004), is perceived as disorder.

According to Edmonds (1995), the problem of perceiving a structure as regular (ordered) or irregular (disordered) lies in the fact that no language of representation of a structure is provided. From the perspective of music, this suggests that the listener, in the absence of an inherent language, has to impose one.

Music is multidimensional, and musical dimensions are never found in isolation but rather constantly interacting (more or less) with each other (Prince *et al.*, 2009b). The perception of these dimensions depends on how they are presented in the musical shape, which can be *vertical*, when the relationships between notes are presented simultaneously (e.g., harmony), or *horizontal* (e.g., melody), when notes are presented in a sequence one after another. A specific example is pitch, which can be presented either vertically (as chord) or horizontally as a sequential presentation of pitches (Loui, 2012). However, vertical and horizontal presentation of music (encompassing all the discussions of what exactly should be considered as vertical or horizontal), when presented together in a musical piece, forms a unit, in which the musical content is deposited (Busch, 1985). These two dimensions are inexorably connected, with horizontal an extension of vertical and vice versa (Williams, 2005). Findings from different studies (Lerdahl and Jackendoff, 1983; Butler and Brown, 1994; Platt and Racine, 1994) suggest that even when only one dimension is presented (e.g., melody), listeners tend to imply structures found in another dimension (e.g., harmony).

According to Sloboda and Parker (1985), each tone in a single melodic line can imply a harmony as a mental model of the underlying structure, with similar findings reported in other studies (Thompson and Cuddy, 1989; Platt and Racine, 1994; Holleran *et al.*, 1995). When melody and harmony are presented together in a musical example (when the vertical and horizontal dimensions are presented together), a harmonic frame is established (Povel and Jansen, 2002), which can have two aspects: a *global aspect* (key and mode) and a *local aspect* defined as a region within the key, which is assigned to a harmony and defined as a function (e.g., tonic, subdominant, dominant, etc.). According to Povel and Jansen (2002), a listener establishes first global and then local aspects, although the processes of the establishment of these two aspects is not well understood, as they are usually conceived as hierarchical (Bharucha, 1987; Tillmann *et al.*, 2000). For AI based analysis of recorded speech, see the work of Piotrowska *et al.* (2019).

The fact that music can be presented vertically and/or horizontally can explain why it is insufficient to capture only one musical dimension (e.g., harmony in a musical piece) when seeking answers what causes higher/lower feelings of (ir)regularity in listeners of music. Findings of Prince *et al.* (2009b) suggest that the perception of a musical dimension alone can differ from the perception of the same musical dimension when interacting with an another dimension. In the latter case, even seemingly small alteration of structure in a particular dimension (e.g., harmony) can affect the the perception of structure in another dimension (e.g., melody). Depending on the stimulus and task properties, the importance of one musical dimension can be magnified (Prince, 2011) depending on the informative value, and a certain dimension, having a greater informative value, is likely to dominate other dimensions (Melara and Algom, 2003; Prince *et al.*, 2009a).

Approaches to measuring musical regularity include evaluation of regularity in music by listeners, which can be very subjective and dependent on long- or short-term exposure and acquisition of formal/informal musical knowledge (Steinbeis *et al.*, 2006; Herbert, 2012). Regardless of the subjective nature of such measurements, they can highlight peculiarities in musical structure, which should be considered when analysing deviations in musical structure (Mihelač and Povh, 2020).

Another example of measuring musical regularity involves mismatch negativity (MMN). Previous studies (Bouwer and Honing, 2012; Bader *et al.*, 2017) indicate that an MMN response depends on the magnitude of the violation (Schröger and Winkler, 1995; Näätänen *et al.*, 2007) and is elicited when the previously established regularity is violated. Because the auditory system is considered predictive in nature (Pearce, 2005; Pearce *et al.*, 2010b; Bouwer and Honing, 2012) and creates expectations based on extracted regularities found in a musical information, any incoming information that does not match a prediction represents an error signal in the form of an MMN (Bendixen *et al.*, 2009).

Another approach to measuring regularity in musical structure is computational simulation of human perception. By enforcing a set of rules, this method examines how well human perception is captured in order to verify theoretical principles and generate quantitative predictions when applying the model to novel circumstances (Plaut, 2000).

We used the latter approach in the present study to examine the regularity of melody in a dataset comprising 160 musical excerpts. We previously studied this dataset in detail (Mihelač and Povh, 2017; 2019; Mihelač *et al.* 2018) and showed that the regularity of the harmonic progression affected user perception while highlighting other musical dimensions. Therefore, we expanded our focus to melody in order to elucidate how regularity in melody and the interplay between melody and harmony affects listener perception.

3. IDyOM

3.1. Modelling musical structure with IDyOM. The computational model IDyOM is based on n-gram models, which are frequently used in statistical language modelling. An n-gram can be defined as a sequence of n symbols, where the sequence itself can include anything (e.g., characters, words, etc.), with an n-gram model which is simply a collection of such sequences. The basic n-gram model used in music is the unigram model, where $n = 1$, and the occurrence of each tone (event) in a sequence is counted.

Because the quantity of $n - 1$ is known as the *order* of the model, this implies that in the case $n = 1$, we have a simple zeroth-order, where each tone is treated as an independent event, i.e., not dependent on the preceding context. For $n = 2$ (a bigram model), the prediction of forthcoming events is governed by a first-order model, and two adjacent tones are considered in the prediction of a forthcoming tone, which depends on the preceding tone.

To capture as much information as possible from musical structure, IDyOM uses a *variable-order* n-gram model. The use of low-order n-gram models, such as unigram or bigram models, might not adequately explain the statistical regularity in musical structure or the effect of context on expectations, and the use of higher order models could prevent the capture of statistical regularity, (see, e.g., Wiggins *et al.*, 2009).

IDyOM learns in an unsupervised manner from the musical structure and generates predictions about forthcoming events in musical sequences (Pearce, 2005). Listeners are sensitive to statistical regularities and irregularities in musical structure, which are progressively internalized during long- or short-term exposure to music and then generalized to new musical examples. IDyOM is capable of simulating both instances, the long- or short-term exposure to music with long- and short-term modelling. Long-term exposure to music is simulated by using the long-term model (LTM), which is trained on a large corpus of music, whereas short-term exposure is simulated using the short-term model (STM), which learns about repeated patterns in a particular musical sequence/piece dynamically (Pearce, 2018). These models have been tested on different tasks and shown to be accurate predictors of melodic expectancy (Pearce and Wiggins, 2006), behaviour, and neural measures (electroencephalography) of melodic expectedness (Pearce *et al.*, 2010c; Agres *et al.*, 2018), as well as accurate identifiers of phrase boundaries (Pearce *et al.*, 2010a; 2010b)

3.2. Viewpoints. IDyOM enables the perception of a sequence of events (e.g., a melody consisting of notes) from different angles. Whenever a sequence of events is given, functions and *viewpoints* are defined to accept initial sub-sequences of a sequence and select a specific feature (e.g., pitch, duration, relationships between tones, etc.) in the sequence (Pearce and Wiggins, 2012). The two crucial dimensions (viewpoints), in which events in a sequence are described in IDyOM, are pitch (`cpitch`) and time (`dur`). IDyOM also offers derived viewpoints, such as `cpint` (the distance between two pitches) and `cpintfref`, with the latter representing how close/far an event in a sequence is from the tonic. The use of this viewpoint is motivated by the fact that the regularities in pitch relative to the tonic influence the melodic structure (Pearce, 2005; 2018; Arthur, 2018).

According to Volk (2016), there is still a gap existing in the understanding between temporal based dimensions and other musical dimensions. Previous studies (Krumhansl, 2000; Justus and Bharucha, 2003) are treating melodic (pitch-based) and temporal (time-based) relations separately, empirically, and theoretically; however, their independence from the perspective of processing has been questioned (Jones and Boltz, 1989; Boltz, 1999; Griffiths *et al.*, 1999), in that melody and rhythm are perceived as a unified dimension by listeners. Therefore, we have decided to link two viewpoints together to create a compound of the viewpoints representing a cross-product of two viewpoints, viewpoint `cpitch`$\otimes$`dur` in the present study, because it remains unclear whether these two dimensions can be considered separately or unified.

The decision to use the viewpoint `cpint` to examine intervals and relationships between two adjacent tones is based on the fact that pitch relationships evoke a particular scale and affect the feeling of stability between scale tones, as all scale tones are not equivalent from the perspective of their importance and are hierarchically organized depending on how distant or closely related they are to the tonic (Tillmann *et al*., 2000; Peretz and Zatorre, 2005).

Another option involves exploring a sequence and its events from multiple viewpoints (viewpoint selection in IDyOM) using flexible views of abstract musical objects. In this case, a hill-climbing procedure is used that combines different viewpoints and identifies the combination of source viewpoints in order to minimize the mean information content of the dataset (Pearce, 2005).

In the present study, we used four viewpoints `cpitch`, `cpint`, `cpintfref`, and `cpitch⊗dur`. These were labelled in our dataset and in the results from Section 4 as `IC_cpitch`, `IC_cpint`, `IC_cpintfref`, and `IC_cpitch⊗dur` for information content and as `E_cpitch`, `E_cpint`, `E_cpintfref`, and `E_cpitch⊗dur` for entropy.

3.3. Entropy and information content. IDyOM uses a complex methodology to compute estimates of probabilities of an event to appear in the sequence that we are considering. It is encapsulated into a lossless data compression algorithm PPM*, an improved version of PPM (Prediction by Partial Match), originally introduced by Cleary and Witten (1984). The classic PPM algorithm, where the maximum context length is a fixed constant, compresses sequences of symbols, one by one, and learns gradually about context-dependent conditional probability distributions (Steinruecken *et al*., 2015). An improved variant of this sequence prediction model, the PPM* algorithm, able to process contexts of unbounded length compared (Cleary *et al*., 1995), has been used in IDyOM, and combined with interpolated smoothing and update exclusion, and long- and short-term models as well (Pearce, 2005). This methodology has been well explored and justified in the last three decades. The reader can find many details in the dissertation by Pearce (2005), in the works of Gold *et al*. (2019) or Pearce (2018) and the references therein. Details about frequency estimating are beyond the scope of this paper, so we take IDyOM as a "black box" that computes estimated probabilities of events based on our data and viewpoints.

The information-theoretic measures used in IDyOM are entropy (H) and information content (IC). Shannon's entropy (Shannon, 1948) is used as baseline theory for quantifying the uncertainty in the prediction of a musical event before it is heard, using specific viewpoint. If X_i is a set of all possible continuations of a given musical event e_i, IDyOM first computes probability p_i that event e_i happens, using the very complex methodology roughly explained in Section 3.1. Then for each $x \in X_i$ it computes probabilities p_x that e_i will continue with x.

Fig. 5. Events in the melody from *The X-Files* (Mark Snow).

Table 1. Probability, entropy, and information content assigned to each event in the sequence for *The X-Files*.

Event	Probability p_i (`cpitch⊗dur`)	$H(e_i)$ (`cpitch⊗dur`)	$IC(e_i)$ (`cpitch⊗dur`)
e_1	0.0470	5.4891	4.4092
e_2	0.0178	4.5753	5.8116
e_3	0.1442	4.6484	2.7932
e_4	0.0173	2.4455	5.8522
e_5	0.0166	3.9282	5.9095
e_6	0.0976	4.3319	3.3555
e_7	0.0110	5.1598	6.5043
e_8	0.0889	5.2377	3.4912
e_9	0.2787	4.5389	1.8427
e_{10}	**0.6076**	2.7924	**0.7187**
e_{11}	**0.5249**	2.3500	**0.9296**
e_{12}	0.0034	2.9635	8.1807
e_{13}	0.0509	5.0373	4.2946
e_{14}	**0.4850**	3.7216	**1.0436**

Once it has p_i and $\{p_x \mid x \in X_i\}$, it computes

$$H(e_i) = -\sum_{x \in X_i} p_x \log_2(p_x), \tag{1}$$

$$IC(e_i) = -\log_2(p_i), \tag{2}$$

The methodology to compute $H(x_i)$ and $IC(x_i)$ is nicely depicted in Fig. 1 from Gold *et al*. (2019).

Figure 5 shows 14 events in the well-known melody for *The X-Files*, a musical excerpt included in the dataset examined in this study and used to examine the probability, information content, and entropy of each event.

Using the viewpoint `cpitch⊗dur`, which examines the feature pitch *and* duration of each event in the sequence, IDyOM computes the corresponding sequence of values for the probability, entropy, and information content (Table 1 and Fig. 6 as `cpitch⊗dur`, depicted as a dashed line).

As shown in Table 1, the events found to be highly expected by the model were events e_{10}, e_{11}, and e_{14} according to their high probability values and low information content. Conversely, events found to be the most unexpected were e_7 and e_{12}, with very low probabilities and high information content.

When analyzing the information content in all four chosen viewpoints in the same musical excerpt, Fig. 6 shows that the highest and lowest information contents were obtained with the viewpoints `cpint` (black full

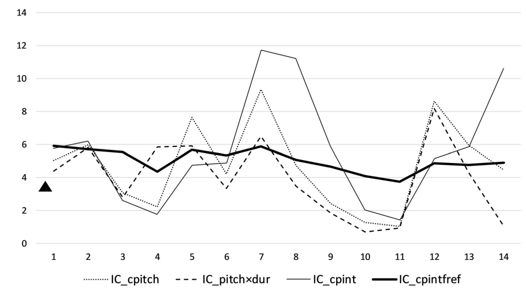

Fig. 6. Visualization of information-content values for each event in the musical excerpt from *The X-Files* and obtained with the four chosen viewpoints: `IC_cpitch`, `IC_cpitch⊗dur`, `IC_cpint`, and `IC_cpintfref`. The black triangle at the beginning of the musical excerpt defines the function in the underlying harmony (a tonic function), which was the same throughout all 14 events.

Table 2. Values of all eight variables representing information content and entropy in the musical excerpt from *The X-Files* and obtained as the arithmetic mean of the corresponding sequence of values for each viewpoint.

Viewpoint	Arithmetic mean
IC_cpitch	4.72288
E_cpitch	3.91749
IC_cpint	5.71109
E_cpint	3.55828
IC_cpintfref	5.04771
E_cpintfref	5.27718
IC_cpitch⊗dur	3.93837
E_cpitch⊗dur	4.08716

line) and `cpitch⊗dur` (black dashed line), respectively. The values reflecting high information content for the viewpoint `cpint` indicate highly unexpected intervals (especially the huge leaps between events e_6 and e_7 and events e_{13} and e_{14}). Conversely, the values for the viewpoint `cpitch⊗dur` showed more expected events from the perspectives of pitch and duration.

The final value of `cpitch⊗dur` for the excerpt from the *The X-Files* was computed as the (arithmetic) mean of the sequence from the last column of Table 1. Similarly, the values of the other seven viewpoints were computed for this excerpt, with the results shown in Table 2. This is a built-in feature of IDyOM.

4. Exploring and predicting regularity with entropy and information content

4.1. Data. The dataset used in our investigations was introduced by Mihelač and Povh (2020) and comprises 160 musical examples, including 141 popular musical pieces and 19 classical music examples, collected and evaluated for their complexity, musical style, and acceptability (Mihelač and Povh, 2020). These musical examples were shortened to musical excerpts from 14 s to 18 s in duration, resulting in 8 to 12 bars. In each musical excerpt, chords in the harmonic progression were located, and entropy (unigram and bigram) was calculated. All of the musical excerpts were labeled as "regular" or "irregular" based on clear criteria which patterns must be present in the excerpt to label it as irregular. When the situation was clear, the label was assigned by the main evaluator, while in ambiguous situations, two additional experts were involved. These values of (ir)regularity were taken as a ground truth in the present research.

For the purpose of this study, we extracted melodies from all these 160 musical excerpts in order to obtain pure monophonic musical excerpts and using only the very first upper lines in all of the melodies. This was done, because some of the melodies in this dataset were written in a polyphonic or chordal manner (see Section 1.3), which would make the analysis of multiple lines in these melodies beyond the scope of this study. The new dataset, with 4692 events in total and an average of 29.33 events per excerpt, comprises all the variables used in the previous study, and additionally it contains eight new variables based on entropy and information content. These eight variables represent mean values of `IC_cpitch, IC_cpint, IC_cpintfref, IC_cpitch⊗dur, E_cpitch, E_cpint, E_cpintfref, and E_cpitch⊗dur` computed for the musical excerpts using computational model IDyOM (for details, see Sections 1.1 and 3.2).

We summarise the distributions of all eight viewpoints in Table 3, which contains the values for min, median, arithmetic mean and max for each viewpoint, separately for regular and irregular excerpts.

4.2. Classification models for regularity. In this subsection, we propose two AI algorithms to predict regularity, as introduced by Mihelač and Povh (2019), which take the information content and entropy variables as inputs and are trained on the dataset from that work (Mihelač and Povh, 2019).

First, the principal component analysis (PCA) was performed on the variables `IC_cpitch, IC_cpint, IC_cpintfref, IC_cpitch⊗dur, E_cpitch, E_cpint, E_cpintfref, and E_cpitch⊗dur`. Figure 7 shows that regular and irregular musical excerpts were linearly separated into a 2-dimensional subspace spanned by the first two principal components, which motivated a search for classification models promoting the linear separation.

Before developing classification models, we explored the relevance of the eight predictive variables using a Mann–Whitney–Wilcoxon test (`R` script `wilcox.test`) to determine which variables showed significantly different distributions on the subgroups of

Table 3. Main characteristics of distributions for all eight viewpoints, separately for regular and irregular excerpts.

	Regular				Irregular			
Viewpoint	min	med	mean	max	min	med	mean	max
IC_cpitch	1.78	3.28	3.17	6.10	1.63	3.48	3.33	5.62
E_cpitch	2.61	3.28	3.26	4.04	2.33	3.45	3.44	4.32
IC_cpint	1.71	3.10	3.11	6.33	1.69	3.02	3.13	5.71
E_cpint	2.49	3.12	3.12	4.01	2.56	3.06	3.05	3.56
IC_cpintfref	3.37	4.39	4.33	5.55	3.66	4.81	4.73	5.74
E_cpintfref	4.08	4.58	4.56	5.00	4.54	5.04	5.01	5.28
IC_cpitch⊗dur	1.59	3.38	3.37	6.10	1.28	3.51	3.50	6.84
E_cpitch⊗dur	2.30	3.39	3.40	4.61	2.30	3.60	3.52	4.42

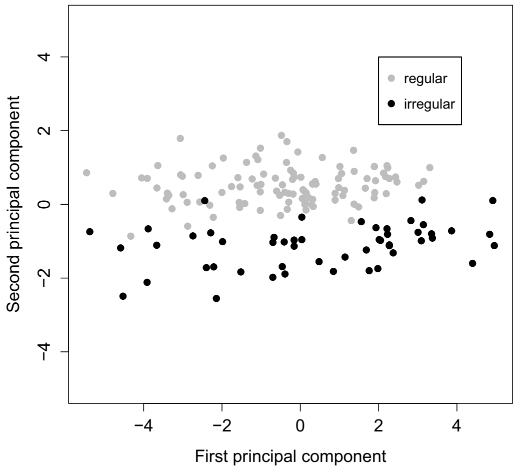

Fig. 7. First two principal components obtained by PCA demonstrated linear separation of the regular and irregular excerpts specifically due to the second principal component.

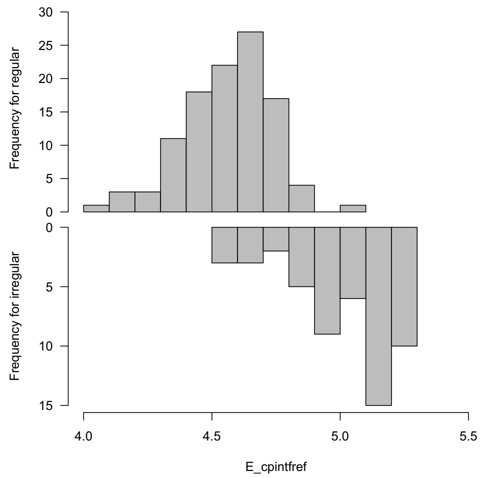

Fig. 8. Distributions of E_cpintfref on the groups of regular and irregular excerpts are very different.

regular and irregular musical excerpts, according to the definition of regularity by Mihelač and Povh (2019). This test was performed because the assumptions requiring the use of two samples t-test were not met. Table 4 shows that the variables E_cpitch, IC_cpintfref, and E_cpintfref had significantly different distributions on the sets of regular and irregular examples. Therefore, they were natural candidates for building blocks of the classification models. Figure 8, which depicts distributions of E_cpintfref on regular and irregular excerpts, additionally supports the decision to use this variable is as a feature in classification models.

We built classification models using subsets of one to eight variables, where in subsets with k variables, the k most relevant variables were used (i.e., the k variables with the lowest p-values in Table 4, which means that the model with one prediction variable included only variable E_cpintfref, the model with two prediction variables included E_cpintfref, IC_cpintfref, etc.).

We used SVM and ANN as classification methods. For the SVM, we used the R library caret and function train with a linear kernel. The training part and evaluation of the models were done using 10-fold cross-validation.

For the ANN, we used the R package neuralnet and the function with the same name. Additionally, we used the logistic activation function. The 10-fold cross-validation was used to build and evaluate the models. Other input settings for neuralnet function were set to default, see the manual (Fritsch *et al.*, 2019) for details about default settings.

Regarding hidden layers, we tested ANN with no hidden layer, with one hidden layer (having 1 to 4 neutrons) and with 2 hidden layers (having 4 neutrons each). The results obtained with zero hidden layers were already very good and increasing the number of layers and the number of neurons did not improve the models significantly (in terms of accuracy and Cohen's Kappa), whereas the computational complexity increased significantly. Therefore, we decided to keep and report the results for ANN with no hidden layers. This actually means that we could replace ANN with logistic regression.

Table 4. Mann–Whitney–Wilcoxon test results indicating that `E_cpitch`, `IC_cpintfref`, and `E_cpintfref` differed significantly between regular and irregular examples.

Viewpoint	p-Value
IC_cpitch	0.275174
E_cpitch	0.003816
IC_cpint	0.971080
E_cpint	0.152142
IC_cpintfref	0.000001
E_cpintfref	0.000000
IC_cpitch⊗dur	0.402339
E_cpitch⊗dur	0.117315

Table 5. Accuracy and Kappa values for SVM and ANN classification models. Each row represents the respective models constructed using the prediction variables (information content and entropy) in order from the most to the least significant (for example, the row labelled "5" corresponds to classification models based on the five most relevant prediction variables).

No. of vars	ACC (SVM)	Kappa (SVM)	ACC (ANN)	Kappa (ANN)
1	0.9179	0.8017	0.8938	0.7636
2	0.9198	0.8030	0.9000	0.7797
3	0.9250	0.8231	0.9313	0.8510
4	0.9374	0.8488	0.9688	0.9340
5	0.9746	0.9397	0.9688	0.9340
6	0.9688	0.9254	0.9812	0.9604
7	0.9691	0.9292	0.9938	0.9868
8	0.9628	0.9164	1.0000	1.0000

We notice that the classes were slightly unbalanced (the class of regular excerpts contained 66.9% of all data). The literature suggests several strategies (Kotsiantis *et al.*, 2006; Krawczyk, 2016) for avoiding bias implied by unbalanced models. For the SVM, we tested three strategies: undersampling, oversampling, and weight adaption of the data instances. We observed only slight changes in the accuracy of the results; therefore, no further explorations were attempted, but we decided to report also Cohen's Kappa which is more adequate for evaluating the models when the classes are not balanced. For ANN, we did not try any strategy to address unbalanced data since the accuracies and Kappas we obtained with the ANN already demonstrated very good predictions. Indeed, Table 5 contains results for ACC and Kappa obtained by SVM and ANN for 1–8 prediction variables.

5. Discussion and conclusions

In this study, we used 160 musical excerpts previously used to evaluate the effect of harmony on musical acceptability (Mihelač, 2017; Mihelač and Povh, 2017; 2020; Mihelač *et al.* 2018; 2019). Given that our focus here was on melody, we extracted from each musical excerpt only the very first upper line in order to obtain 160 monophonic musical excerpts, which were then used to evaluate eight variables, including the pitch, interval, scale degree, and duration of the information content (`IC_cpitch, IC_cpint, IC_cpintfref, and IC_cpitch⊗dur`) and the entropy (`E_cpitch, E_cpint, E_cpintfref, and E_cpitch⊗dur`) by adding them to the original dataset.

To identify irregularities in 53 of the 160 musical excerpts, which was detected during the previous study and were taken as ground truth in the present study, we analyzed the relevance of the new features using the Mann–Whitney–Wilcoxon test, revealing that three variables (`E_cpitch,IC_cpintfref, and E_cpintfref`) differed significantly between regular and irregular musical excerpts.

The significant difference between regular and irregular musical excerpts demonstrated by `IC_cpintfref and E_cpintfref`, suggests the presence of implied harmonies, and confirms the salience of implied harmonies in the perception of music in homophonic music (Sloboda and Parker, 1985; Thompson and Cuddy, 1989; Platt and Racine, 1994; Holleran *et al.*, 1995).

Applying the concepts of "global" and "local" establishment of the harmonic frame from Povel and Jansen (2002) to previous and new data revealed that while listening to a particular monophonic musical excerpt, listeners first generate a key and mode, after which implied harmonies are "created" for each note. In some cases, these implied harmonies do not "fit" in the existing harmonic framework (according to the rules of the harmonic syntax), which is (re)created when the same melody is combined with its underlying harmony.

Therefore, when "horizontal" (melody) and "vertical" (harmony) musical content are presented together, a "fusion" of different tones occurs (Parncutt, 1989; Huron, 2001) to generate different harmonies. These harmonies can be emphasized depending on the information value either in melody or harmony, which could explain the feeling of higher complexity in musical excerpts with a simple harmonic progression (e.g., T-D-T), as the focus is placed on the melody and its harmonies. This agrees with previous findings reported by Melara and Algom (2003) or Prince *et al.* (2009a).

The significance of the variable `E_cpitch` can be explained by the higher degree of diversity in pitches and the higher number of non-chordal tones found in irregular musical excerpts. Specifically, non-chordal tones appear to affect the listener enjoyment of music, as well as the expectation for forthcoming events, according to our previous results (Mihelač *et al.*, 2018; Mihelač and Povh, 2020). The successful harmonic analysis of a musical piece by a listener is clearly dependent on how successfully the non-chordal tones are resolved and assigned to a harmony and how the tones distributed in a melody are perceived as an implied harmony, which agree with findings from a previous study (Povel and Jansen, 2002).

In our previous work (Mihelač and Povh, 2020), we identified 10 peculiarities found in irregular musical excerpts and based on the musical expertise of three experts according to their evaluation and perception of complexity. In the present study, we used two classification methods (SVMs and ANNs), with both algorithms using 10-fold cross-validation to confirm a high level of accuracy (>97%) in predicting (ir)regularity in our dataset. These results indicated that expert-based detection of (ir)regularity in musical structure can be replaced by AI algorithms.

Because only eight variables were included in the analysis of regularity of musical structure in this study, future work should focus on analysis of additional viewpoints (variables) in IDyOM. This was previously found useful in the perception of musical structure and prediction of forthcoming events (Pearce, 2005), with additional study potentially offering deeper insight into the (ir)regularity of musical structure. Additionally, the approaches used in the present study to analyse (ir)regularity could be applied to other datasets, especially to those comprising non-complex/less complex musical examples (e.g., children's songs, children's folk songs, folk songs, etc.), and used as an additional clarification of the listener's acceptance/rejection of musical pieces/genre.

Acknowledgment

This research work was supported by Šolski Center Novo Mesto, Slovenia, and by the Slovenian Research Agency (ARRS) through the research project J1-8155 and the research program P2-0256.

References

Agres, K.R., Abdallah, S. and Pearce, M.T. (2018). Information-theoretic properties of auditory sequences dynamically influence expectation and memory, *Cognitive Science* **42**(1): 43–76.

Arthur, C. (2018). A perceptual study of scale-degree qualia in context, *Music Perception* **35**(3): 295–314.

Bader, M., Schröger, E. and Grimm, S. (2017). How regularity representations of short sound patterns that are based on relative or absolute pitch information establish over time: An EEG study, *PlosONE* **12**(5): e0176981.

Bendixen, A., Schröger, E. and Winkler, I. (2009). I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system, *The Journal of Neuroscience* **29**(26): 8447–8451.

Benward, B. and Saker, M. (2008). *Music in Theory and Practice*, William Glass, New York, NY.

Bharucha, J.J. (1987). Music cognition and perceptual facilitation: A connectionist frame-work, *Music Perception* **5**(1): 1–30.

Boltz, M.G. (1999). The processing of melodic and temporal information: independent or unified dimensions?, *Journal of New Music Research* **28**: 67–79.

Bouwer, F. and Honing, H. (2012). Rhythmic regularity revisited: Is beat induction indeed pre-attentive?, *Proceedings of the 12th International Conference of the European Society for the Cognitive Sciences of Music, Thessaloniki, Greece*, pp. 122–127.

Burns, E.M. (1999). Intervals, scales, and tuning, *in* D. Deutsch (Ed.), *The Psychology of Music*, Academic Press, New York, NY, pp. 215–264.

Busch, R. (1985). On the horizontal and vertical presentation of musical ideas and on musical space (I), *Tempo* (154): 2–10.

Butler, D. and Brown, H. (1994). Describing the mental representation of tonality in music, *in* R. Aiello and J.A. Sloboda (Eds), *Musical Perceptions*, Oxford University Press, New York, NY, pp. 191–212.

Cleary, J.G., Teahan, W.J. and Witten, I.H. (1995). Unbounded length contexts for PPM, *Proceedings of the Data Compression Conference, DCC'95, Snowbird, UT, USA*, pp. 52–61.

Cleary, J.G. and Witten, I. (1984). Data compression using adaptive coding and partial string matching, *IEEE Transactions on Communications* **32**(4): 396–402.

Dahlhaus, C. (2014). *Studies on the Origin of Harmonic Tonality*, Princeton University Press, Princeton, NJ.

Edmonds, B. (1995). What is complexity? The philosophy of complexity per se with application to some examples in evolution, *The Evolution of Complexity, Brussels, Belgium*.

Feldman, J. (1997). Regularity-based perceptual grouping, *Computational Intelligence* **13**(4): 582–623.

Finnas, L. (1989). How can musical preferences be modified? A research review, *Bulletin of the Council for Research in Music Education* **102**: 1–58.

Fritsch, S., Guenther, F. and Guenther, M.F. (2019). Package 'neuralnet', `https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf`.

Gold, B.P., Pearce, M.T., Mas-Herrero, E., Dagher, A. and Zatorre, R.J. (2019). Predictability and uncertainty in the pleasure of music: A reward for learning?, *Journal of Neuroscience* **39**(47): 9397–9409.

Grassberger, P. (2004). Problems in quantifying self-organized complexity, *Helvetica Physica Acta* **62**(5): 498–511.

Griffiths, T.D., Johnsrude, I., Dean, J.L. and Green, G.G. (1999). A common neural substrate for the analysis of pitch and duration pattern in segmented sound?, *NeuroReport* **10**(18): 3825–3830.

Herbert, R. (2012). Young people's use and subjective experience of music outside school, *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, Thessaloniki, Greece*, pp. 423–430.

Holleran, S., Jones, M.R. and Butler, D. (1995). Perceiving implied harmony: The influence of melodic and harmonic context, *Journal of Experimental Psychology: LMC* **21**(3): 737–753.

Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles, *Music Perception* **11**(1): 1–64.

Jackendoff, R. (2009). Parallels and nonparallels between language and music, *Music Perception* **26**(3): 195–204.

Jones, M.R. and Boltz, M. (1989). Dynamic attending and responses to time, *Psychology Revue* **96**(3): 459–491.

Justus, T. and Bharucha, J. (2003). Music perception and cognition, *in* A. Yantis and H. Pasler (Eds), *Stevens Handbook of Experimental Psychology, Volume I: Sensation and Perception*, Wiley, New York, NY, pp. 453–492.

Kotsiantis, S., Kanellopoulos, D. and Pintelas, P.E. (2006). Handling imbalanced datasets: A review, *GESTS International Transactions on Computer Science and Engineering* **30**(1): 25–36.

Kramer, J.D. (1988). *The Time of Music: New Meanings, New Temporalities, New Listening Strategies*, Schirmer Books, New York, NY.

Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions, *Progress in Artificial Intelligence* **5**(4): 221–232.

Krumhansl, C.L. (2000). Rhythm and pitch in music cognition, *Psychological Bulletin* **126**(1): 159–179.

Krumhansl, C.L. (2004). The cognition of tonality—As we know it today, *Journal of New Music Research* **33**(3): 253–268.

Lerdahl, F. and Jackendoff, R. (1983). An overview of hierarchical structure in music, *Music Perception* **1**(2): 229–252.

Loui, P. (2012). Learning and liking of melody and harmony: Further studies in artificial grammar learning, *Topics in Cognitive Science* **4**(4): 554–567.

Manjunath, B.S., Wu, P., Newsam, S. and Shin, H.D. (2000). A texture descriptor for browsing and similarity retrieval, *Signal Processing: Image Communication* **16**(1–2): 33–43.

Melara, R.D. and Algom, D. (2003). Driven by information: A tectonic theory of stroop effects, *Psychological Review* **110**(3): 422–471.

Meyer, L.B. (1957). Meaning in music and information theory, *Journal of Aesthetics and Art Criticism* **15**(4): 412–424.

Mihelač, L. (2017). *Napovedovanje slušne sprejemljivosti na osnovi entropije harmonije*, Master's thesis, School Center, Novo Mesto.

Mihelač, L. and Povh, J. (2017). Predicting the acceptability of music with entropy of harmony, *14th International Symposium on Operations Research in Slovenia, SOR'17, Bled, Slovenia,* pp. 371–375.

Mihelač, L. and Povh, J. (2020). The impact of the complexity of harmony on the acceptability of music, *ACM Transactions on Applied Perception* **17**(1): 1–27.

Mihelač, L. and Povh, J. (2019). The impact of harmony on the perception of music, *15th International Symposium on Operations Research in Slovenia SOR'19, Bled, Slovenia*, pp. 360–365.

Mihelač, L., Wiggins, A.G., Lavrač, N. and Povh, J. (2018). Entropy and acceptability: Information dynamics and music acceptance, *Proceedings of ICMPC15/ESCOM10, Graz, Austria*, pp. 313–317.

Näätänen, R., Paavilainen, P., Rinne, T. and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review, *Clinical Neurophysiology* **118**(12): 2544–2590.

Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication–Realisation Model*, University of Chicago Press, Chicago, IL.

Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach*, Springer, Berlin.

Pauly, M., Mitra, N.J., Wallner, J., Pottmann, H. and Guibas, L.J. (2008). Discovering structural regularity in 3D geometry, *ACM Transactions on Graphics* **27**(3), Article no. 43.

Pearce, M.T. (2005). *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition*, PhD thesis, City University, London.

Pearce, M.T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation, *Annals of the New York Academy of Sciences* **1423**(1): 378–395.

Pearce, M.T., Müllensiefen, D. and Wiggins, G.A. (2010a). Melodic grouping in music information retrieval: New methods and applications, *in* Z.W. Raś and A.A. Wieczorkowska (Eds), *Advances in Music Information Retrieval*, Springer, Berlin, pp. 364–388.

Pearce, M.T., Müllensiefen, D. and Wiggins, G.A. (2010b). The role of expectation and probabilistic learning in auditory boundary perception: A model comparison, *Perception* **39**(10): 1365–1389.

Pearce, M.T., Ruiz, M.H., Kapasi, S., Wiggins, G.A. and Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation, *NeuroImage* **50**(1): 302–313.

Pearce, M.T. and Wiggins, G.A. (2012). Auditory expectation: The information dynamics of music perception and cognition, *Topics in Cognitive Science* **4**(4): 625–652.

Pearce, M. and Wiggins, G. (2006). The information dynamics of melodic boundary detection, *Proceedings of the 9th International Conference on Music Perception and Cognition, Bologna, Italy*, pp. 860–865.

Peretz, I. and Zatorre, R.J. (2005). Brain organization for music processing, *Annual Review of Psychology* **56**: 89–114.

Piotrowska, M., Korvel, G., Kostek, B., Ciszewski, T. and Czyżewski, A. (2019). Machine learning-based analysis of English lateral allophones, *International Journal of Applied Mathematics and Computer Science* **29**(2): 393–405, DOI: 10.2478/amcs-2019-0029.

Platt, J.R. and Racine, R.J. (1994). Detection of implied harmony changes in triadic melodies, *Music Perception* **11**(3): 243–264.

Plaut, D.C. (2000). Methodologies for the computer modeling of human cognitive processes, *in* F. Boller *et al.* (Eds), *Handbook of Neuropsychology*, Elsevier, Amsterdam, pp. 259–267.

Pole, W. (2014). *The Philosophy of Music*, Routledge, Taylor & Francis, Milton Park, Abingdon.

Povel, D.-J. and Jansen, E. (2002). Harmonic factors in the perception of tonal melodies, *Music Perception* **20**(1): 51–85.

Prince, J.B. (2011). The integration of stimulus dimensions in the perception of music, *Quarterly Journal of Experimental Psychology* **64**(11): 2125–2152.

Prince, J.B., Schmuckler, M.A. and Thompson, W.F. (2009a). The effect of task and pitch structure on pitch-time interactions in music, *Memory and Cognition* **37**(3): 368–381.

Prince, J.B., Thompson, W.F. and Schmuckler, M.A. (2009b). Pitch and time, tonality and meter: How do musical dimensions combine?, *Journal of Experimental Psychology Human Perception & Performance* **35**(5): 1598–1617.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony, *Journal of Mathematics and Music* **5**(1): 35–53.

Rohrmeier, M. and Pearce, M.T. (2018). Musical syntax I: Theoretical perspectives, *in* R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, Springer, Berlin/Heidelberg, pp. 473–486.

Schröger, E. and Winkler, I. (1995). Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. viewpoint systems for music prediction, *Neuroscience Letters* **193**(3): 185–188.

Shannon, C.E. (1948). A mathematical theory of communication, *Bell System Technical Journal* **27**(3): 379–423.

Sloboda, J.A. and Parker, D.H.H. (1985). Immediate recall of melodies, *in* P. Howell *et al.* (Eds), *Musical Structure and Cognition*, Academic Press, London, pp. 143–167.

Solomon, J.W. (2019). *Music Theory Essentials—A Streamlined Approach to Fundamentals, Tonal Harmony, and Post-Tonal Materials*, Taylor & Francis, Milton Park, Abingdon.

Steinbeis, N., Koelsch, S. and Sloboda, J.A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses, *Journal of Cognitive Neuroscience* **18**(8): 1380–1393.

Steinruecken, C., Ghahramani, Z. and MacKay, D. (2015). Improving ppm with dynamic parameter updates, *Data Compression Conference 2015, Snowbrid, UT, USA*, pp. 193–202.

Thompson, W.F. and Cuddy, L.L. (1989). Sensitivity to key change in chorale sequences: A comparison of single voices and four-voice harmony, *Music Perception* **7**(2): 151–168.

Tillmann, B., Bharucha, J.J. and Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach, *Psychological Review* **107**(4): 885–913.

Volk, A. (2016). Computational music structure analysis: A computational enterprise into time in music, *in* M. Müller *et al.* (Eds), *Computational Music Structure Analysis*, Vol. 6, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Leibniz, p. 159.

Wiggins, G.A., Pearce, M.T. and Müllensiefen, D. (2009). Computational modeling of music cognition and musical creativity, *in* R.T. Dean (Ed.), *Oxford Handbook of Computer Music and Digital Sound Culture*, Oxford University Press, Oxford.

Williams, L.R. (2005). Effect of music training and musical complexity on focus of attention to melody or harmony, *Journal of Research in Music Education* **53**(3): 210–221.

Lorena Mihelač currently works as a VET professor and a supervisor in the SciDrom Scientific Lab at the School Center Novo Mesto, Slovenia. She holds a BA from the University of Music in Zagreb, Croatia, an MA from the Academy of Arts in Novi Sad, Serbia, as well as a BSc and an MSc from the Faculty of Information Studies of the School Center Novo Mesto, Slovenia. She also holds a PhD from the Postgraduate School ZRC SAZU, Ljubljana, Slovenia. Now she is finishing her second PhD at Jožef Stefan International Postgraduate School (Information and Communication Technologies) in Ljubljana. Her research covers different areas, such as music education, music psychology, music therapy, and lately modelling of human music perception. She has published many educational textbooks, several chapters in monographs, journal papers, and numerous works in conference proceedings.

Janez Povh is an associate professor and a senior researcher at the Faculty of Mechanical Engineering, University of Ljubljana. He holds BSc, MSc and PhD degrees in mathematics from the University of Ljubljana. His research activities vary from pure mathematical optimization topics, like developing new methods for (non-commutative) polynomial optimization problems, to more applied areas such as applications of mathematical optimization methods and tools in data science and complex networks analysis. He has published 35 journal papers (17 in top mathematical optimization journals), numerous works in conference proceedings, several book chapters and a book with Springer.

Received: 7 January 2020
Revised: 4 May 2020
Accepted: 20 June 2020