BACHELOR THESIS

New rules for the Melody Generator II
Repeating rhythm and varying harmony

Student:
Marieke SWEERS
s3046907

Supervisors:
Phd. Makiko SADAKATA
Dr. Louis G. VUURPIJL

September 3, 2012

New rules for the Melody Generator II
Repeating rhythm and varying harmony

Marieke Sweers
Phd. Makiko Sadakata
Dr. Louis G. Vuurpijl
Radboud University Nijmegen (Social science, Artificial Intelligence)

Abstract

New rules were sought for the Melody Generator II, a knowledge based system that composes music. The system lacked good rules for continuation of a melody, so several features were selected to investigate how the melodies could continue. The features - harmony, contour and rhythm - were manipulated such that they could be repeated or varied. We found that it was best to repeat the rhythm over the course of the melody and to vary the harmony when the contour was repeated.

Keywords: Computer generated music, Knowledge Based System, Continuation of a Melody, Melody Generator II, Repetition, Variation

Introduction

Can computers compose music? Researchers have been interested in this question for over 50 years. Multiple ways to make a computer compose music have been introduced. These include systems which learn from examples (statistical learning), evolutionary methods and knowledge based systems (Papadopoulos & Wiggins, 1999). A nice example of learning systems is Cope's work, the system extracts information from existing pieces of music to create new music (Cope, 1996). Evolutionary methods include the work of McIntyre (1994) and an example of a knowledge based system (KBS) is the program introduced by Povel (2010).

It has been said that a KBS is not suitable for creating music. A problem with these systems is that they need rules from music theory, but music theory is not formal enough (Delgado, Fajardo, & Molina-Solana, 2009). However, this lack of a formal description in music theory can be complemented with other information coming from statistical information (for example from work like Ponsford, Wiggins, and Mellish (1999) did) and with rules that are extracted from experiments as done here.

The KBS made by Povel (2010), the Melody Generator II, is based largely on theories by Deutsch and Feroe (1981) and Lerdahl and Jackendoff (1996) and it actually works quite well, although it is not as good as a human composer.

A knowledge based system: Melody Generator II

The current study searches new rules for a KBS: the Melody Generator II (MG II). MG II already includes many rules that (Povel, 2010) based among others on theories from work by Lerdahl and Jackendoff (1996), Longuet-Higgins and Lee (1982), Deutsch and Feroe (1981) and Piston and Devoto (1987). The rules are about the features of a melody, for example rhythm, harmony and contour. MG II contains four models which make melodies in different ways: Attraction-based, Scale-based, Chord-based and the basic model (see Povel (2010) for more details). These models use different rules for the features to make the melodies. In this experiment the Chord-based model was used; it makes a melody in the following way.

A melody is made by applying rules from the different features one at a time, starting with rhythm. The rhythm is determined by the meter, the number of syncopations and the rhythmical density. Rules state that notes are placed at metrically strong positions, for example at the beginning of a measure. By changing the number of syncopations by hand, melodies with less metrical stability can be made. Then harmony is set. It can be chosen from a range of built in progressions, for example I-IV-V-I. The contour of the melody can then be chosen from a predefined set, including ascending and sinusoid. Only the direction of the intervals is set here, not of the size of the intervals. After these three are determined, the contour is filled in with skeleton tones and ornamental tones. First the skeleton tones are placed, then the ornamental tones are filled in. Skeleton notes are mostly placed on metrically strong positions and are all chord tones as the name of the model suggests. Ornamental tones can be any tone, they may be chord tones or non-chord tones.

By combining these different rules, the system creates a good beginning of a music piece, a theme. However, it decides little about how the theme should continue. The system offers full possibilities to vary or repeat all feature parameter settings for the following part in a melody. Even so, this means that there is no specific knowledge of how these parameters should be set to continue. Maybe it is good to repeat all features, or it is better to repeat some but to vary others. A rule is needed to describe this continuation. The current study will look into the continuation of the melody: which features should be repeated and which features should be varied. Three features (harmony, contour and rhythm) have been selected for this purpose.

Features of Repetition

There are various features of a piece of music that can be repeated; in this study the focus lies on harmony, contour and rhythm. The definition for contour is slightly different from the definition Povel uses. Every feature will have two levels: repeated and varied.

Rhythm is set in a meter (e.g. $\frac{3}{4}$ or $\frac{4}{4}$) which divides time in periods called measures that are repeated and can be divided into small periods. The various periods have different metrical weights, the first period of a measure has the strongest metrical weight. The higher the weight, the more accentuated a note will be. A rhythm is a pattern of note durations within a meter. It may contain syncopation: a delay in the rhythm or an accent where there normally would not be one.

The *harmony* of a melody is the succession of its chords, mostly groups of three tones, that are played to accompany and guide the melody. Chords are built within a scale, for example C major where the C chord is the tonic, the first and most important chord. The dominant, the G chord, is the fifth chord and is second in rank of importance. The different chords that are used stand in relation to one another, independent of the scale that is used.

The *contour* is the sequence of relative pitch heights, or intervals, of the successive tones in a melody. It contains information about the direction of the interval and the size. Take for example Frère Jacques: the first interval (between the first and second note) goes up one tone, from a C to a D; the second up one tone as well from D to E; the third interval goes down two steps, from E to C. Mostly, more important (emphasized) notes in a melody have the same tones as the tones of the underlying chord.

The three features can all be repeated or varied during a piece of music independently. It is clear that rhythm is completely independent of the contour and harmony: the duration of a note can change without changing its pitch height. Harmony and contour are independent as well. The relative pitch height of notes can be varied while repeating the chords and the chords can be varied without changing the intervals. If the chord moves up or down a few steps the melody can simply be moved up or down the same number steps, leaving the contour intact.

Since every feature can be repeated or varied independently, eight different combinations of the features were made, see table 1. The conditions are coded by three letters ('h' for harmony, 'c' for contour, 'r' for rhythm) with a corresponding number, zero meaning repeated, one varied. A melody was made by combining four units (small pieces of a melody), the first unit was the standard and the next units were derived from the first. Example melodies for every condition can be found in figure 5. As you can see, when a feature was varied, it varied three times, from unit one to two, from unit two to three and from three to four.

Hypothesis

The question this research will answer is which features should be repeated and which should be varied to make the music sound best. The answer to this question will give insight into what is important to be repeated in a melody to make the music sound as a whole. Rhythm is expected not to influence how good a melody will sound, since we did not hear a difference between conditions with rhythm repetition and variation. Condition 3, h1c0r0, is expected to sound the best since there will be a melody line (the contour) that will lead the listener through the music, but there will be some variation in the form of a different harmony. Since rhythm is not expected to make a large difference condition 5 (h1c0r1) will sound almost the same and in such is ranked directly below as second. Furthermore, condition 1 (h0c0r0) will probably sound good as well, to us it sounded like a antecedent and consequent phrase. It is ranked third and condition 2 (h0c0r1) is ranked fourth. Condition 8, where every feature is varied (h1c1r1) is thought to sound worst: there is no point of reference for people to hear where the melody is going. It will sound like random notes tied together and is given the lowest rank, with condition 7 (h1c1r0) just above ranking seventh. About conditions 4 and 6 no specific predictions were made. They were placed in-between the conditions that were predicted to sound worst and those that were predicted to sound best.

To test the hypotheses two experiments were conducted where people were asked to listen to melodies with the different combinations of the features. First they were asked to rate melodies how good they sounded. The second experiment was done to understand how similar the successive parts of music of the different conditions sounded to participants.

Experiment 1

In order to investigate which combination of repetitions sounded best, people were asked to rate how human-made a given melody sounded (naturalness) and to give their preference for this melody.

Method

Participants A total of 78 volunteers completed this experiment, of that total 76 (17-80 years old, 42 women, 34 men) completed experiment 2 as well. For our analysis, the data of the two participants who completed only experiment one were left out.

Materials The materials consisted of MIDI melodies generated with the Chord-Based model of the Melody Generator II. A melody consisted of 4 units, each 2 bars in length, giving a total of 8 bars per melody. The features (harmony, contour and rhythm) were manipulated as described below. They were varied or repeated at

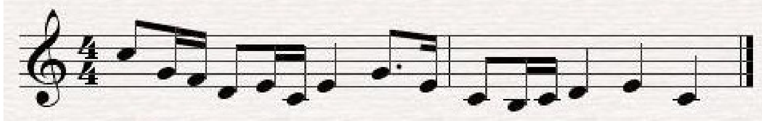
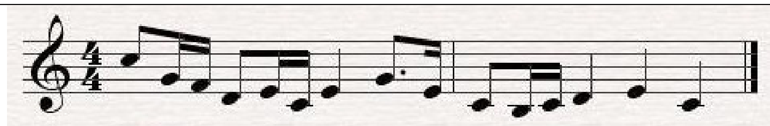
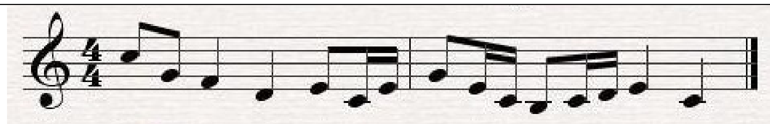
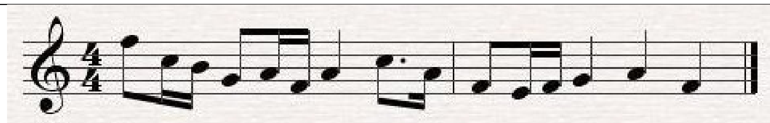
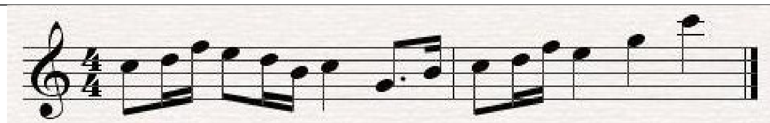
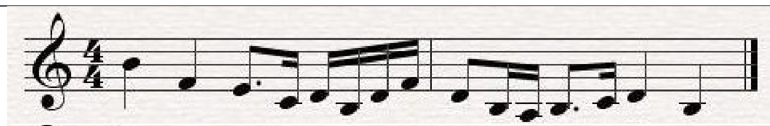
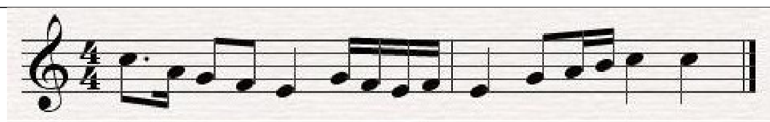
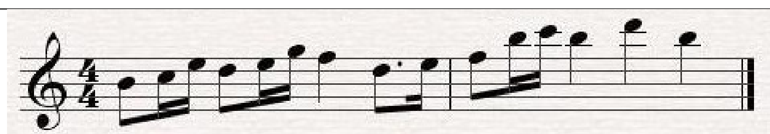
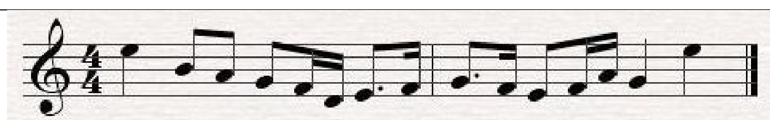
Standard			Hypothesis (Rank in natural-ness)
1	h0c0r0		3
2	h0c0r1		4
3	h1c0r0		1
4	h0c1r0		5
5	h1c0r1		2
6	h0c1r1		6
7	h1c1r0		7
8	h1c1r1		8

Table 1: Example units per condition. The eight units are all derived from the standard. The coding of every condition can be read as follows: the three letters (h, c and r) are abbreviations for the features (harmony, contour and rhythm). The number behind the letter belongs to it and can be zero or one, coding respectively for repeated or for varied. For example, the first condition contains only zero's, everything is repeated. As you can see, the unit is an exact copy of the standard melody. The sixth condition (coding: h1c0r1) has a varied harmony, a repeated contour and a varied rhythm. Every unit that has a varied harmony is transposed a random number of steps. Compare unit five with the standard: unit five was transposed up four steps within the key. Every unit with a varied contour has different intervallic distances between successive notes. Take for example the first interval of unit three: the interval between the first and the second note is a distance of one step up, while the standard unit has an interval of four steps down. A unit with a varied rhythm contains the same notes as the standard unit, only the length of the notes is changed. Compare unit two with the standard: All the notes of the standard unit are present in unit two, only the duration is changed. In the right column our hypothesis is shown. The condition that was hypothesized to sound best or most natural, h1c0r0, was giving rank 1. The condition that was thought to sound worst was h1c1r1.

the second, third and fourth units and this resulted in 8 conditions.

For example, in condition one, the first unit was simply repeated four times without any variations. In condition six, the harmony and rhythm were varied independently of one another in every unit, while the contour was repeated in every unit. Example melodies are described in figure 5.

The *rhythm* was determined by the number of notes per bar and the strength of syncopation. The number of notes was chosen randomly using a range of note types (16^{th} notes to quarter notes), fitting in bars with a $\frac{4}{4}$ meter. The strength of syncopation was also randomly chosen, allowing a maximum of two syncopations within a bar. When the rhythm was varied, both the amount of syncopation and then the placement of the notes was changed.

The *harmony* of the first unit was chosen randomly from three basic progressions: I-V, I-I and V-I of a major key. A base note of chords accompanied the melody to ensure that participants could, implicitly or explicitly, hear the harmonic progression. If the harmony of a melody was varied, all notes of a unit, including the base note, were transposed up or down an equal amount of steps within the used key, thereby changing the harmony.

The *contour* was created by setting the direction of every step (up or down) from one note to the next and then selecting the exact height of the notes. The steps were set to random, no specific pattern (e.g. U-shaped, where the first steps of the melody are down and the last steps up) was used. When contour was varied, the contour was set again.

A total of 192 melodies was created, 24 for each condition (2 melodies per key, using 12 keys). Different keys were included to assure that any effect was not due to a specific range of pitch height. Employing different keys was assumed to have no influence on the ratings.

Procedure Participants were invited to go to the experiment website. First their age, gender and native language were asked and the sound system could be adjusted. After that the following instructions were provided:

For each melody, you are asked to fill in two questions: how human-like the melody was and how much you liked it. Both questions can be rated on an eight-point scale. The melody can only be played once, so please make sure that you are ready to listen before pressing the "Play Melody" button.

Participants were then given 2 practice trials, each containing a melody and its rating form. This was followed by the experiment session that consisted of 16 trials (2 trials per condition), that were randomly presented.

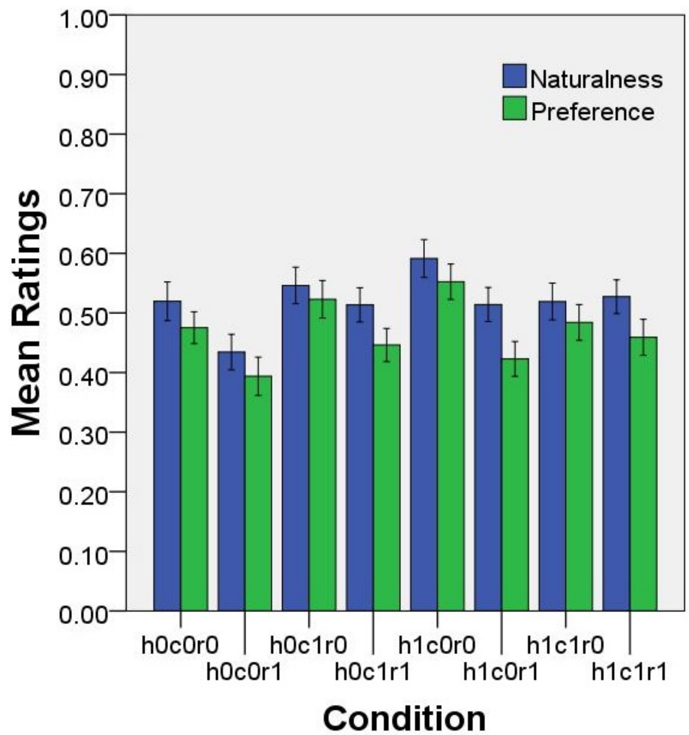


Figure 1: Mean ratings of Naturalness and Preference towards a melody per condition. Conditions are ordered according to amount of features varied. Features are ordered according to importance of results from experiment one: rhythm is displayed first, then harmony then contour.

Results

The two ratings (naturalness and preference) were first normalized according to individuals response range. Figure 1 presents average normalized ratings for 8 conditions in order of varied features. It shows that both ratings varied in between 0.4 - 0.6. Three-way repeated measures analyses of variance (ANOVA) were performed with harmony, contour and rhythm (repeated vs. varied) independently for naturalness and for preference ratings. See table 2 for summary statistics. A significant correlation between preference and naturalness rating was observed, r_s=.56, $p<.01$. Further analyses are given in the following sections for each rating.

Naturalness The ANOVA revealed a significant main effect of rhythm with a significant interaction between rhythm and contour. The interaction between harmony and contour was also significant. Figure 2 illustrates these two interactions. The top panel shows that participants tended to rate melodies more natural when rhythms were repeated, but this was significant only when the contour was repeated. The bottom panel shows that participants tended to rate melodies more natural when harmony was varied and the effect was significant

Effects	Naturalness	Preference
Rhythm	*F(1,77)=5.12, p<.05*	*F(1,77)=5.12, p<.001*
Harmony	F(1,77)=3.30, p=.07	F(1,77)=1.24, p=.27
Contour	F(1,77)=0.60, p=.42	F(1,77)=0.50, p=.48
Rhythm * Harmony	F(1,77)=0.61, p=.40	F(1,77)=0.00, p=.99
Rhythm * Contour	*F(1,77)=4.24, p<.05*	F(1,77)=2.12, p=.15
Harmony * Contour	*F(1,77)=4.05, p<.05*	F(1,77)=2.70, p=.11
Rhythm * Harmony * Contour	F(1,77)=0.11, p=.74	F(1,77)=2.20, p=.14

Table 2: Results from the analyses of Naturalness and of Preference. Significant effects are printed in italic.

only when the contour was repeated. When contour was varied, both effects disappeared.

Preference The ANOVA revealed a significant main effect of rhythm. Participants tended to prefer melodies with a repeated rhythm over that with varied rhythms (see figure 3).

Results indicated that our manipulation indeed influenced the ratings. Especially the effect of rhythm was somewhat consistent for both ratings. Our next question is to what extent the different forms of variation (harmony, rhythm and contour) influence a melody. It might be that varying one feature changes a melody much more drastically than varying another feature. We would like to know how repeated and varied participants perceived the melodies of the different conditions.

Results indicated that our manipulation indeed influenced the ratings. Especially the effect of rhythm was somewhat consistent for both ratings. Our next question is if our manipulation influenced how repeated the melodies sounded. If the melodies are perceived as different in similarity, there is extra evidence that it did influence the repetition.

Experiment 2

Participants were presented with two melodies and asked to give their similarity rating.

Method

Participants The subjects from Experiment 1 that were used in analysis took part in this experiment as well.

Materials The materials consisted of MIDI melodies generated by the Chord-Based model of the Melody Generator II. Pairs of melodies were created; every melody had a length of two bars to make it easy for the listener

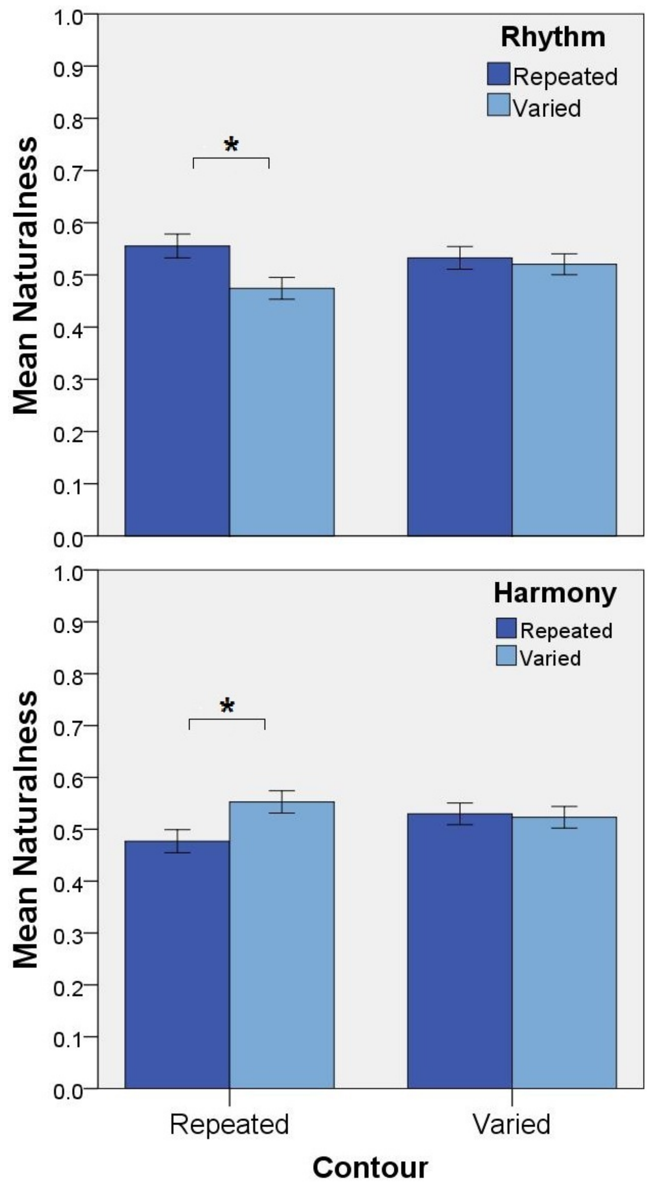


Figure 2: Mean naturalness ratings. Results for rhythm (top panel) and harmony (bottom panel) are shown seperately for the contour conditions. The error bars indicate the standard errors.

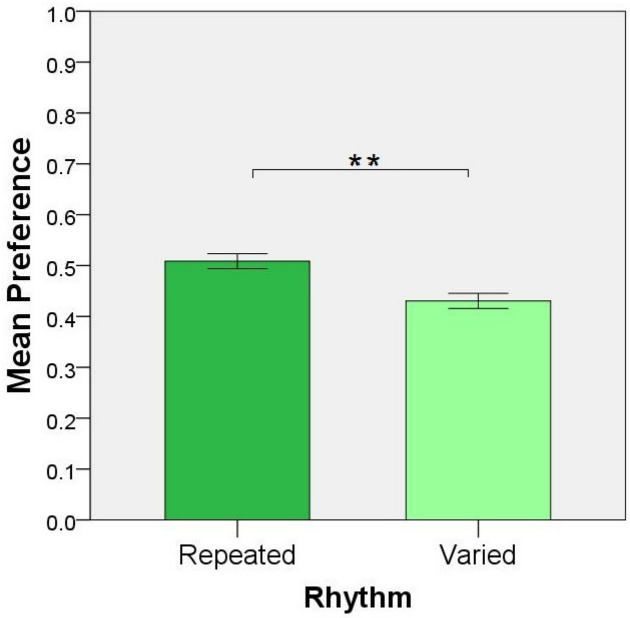


Figure 3: Mean preference ratings for the two rhythm conditions.

to remember the melodies when rating them. The first of a pair was a standard melody and the second a comparison melody. The three features (rhythm, contour and harmony) of the standard melody were repeated or varied in order to create 8 comparison melodies (one per condition). Twelve standard melodies were created (one melody per key) and each had 8 comparison melodies, resulting in 96 melody pairs.

Procedure Participants were given the following instructions:

You will listen to pairs of melodies. After hearing both melodies of each pair, you are asked to rate their similarity on a four-point scale.

The experiment session consisted of eight trials, one per condition. The trials were randomly ordered.

Results

Figure 4 shows average similarity ratings as a function of conditions, ordered according to the number of varied features from left to right. There was a significant correlation between the number of varied features and the similarity rating, r = -.61, p<.01. Participants tended to rate melody pairs more similar when there were more shared features between a given pair. For every pair of conditions it was tested if the difference in ratings was significant and how large this effects was. The comparisons are shown in table 3. Only effect sizes of pairs with significantly different ratings were shown; α was set to 0.00178 to adjust for the high amount of comparisons. As you can see in the top row, condition h0c0r0 differed significantly from every other condition and this effect was very large, r was over .5 for every combination. Furthermore, conditions with one varied feature differed significantly from conditions with more varied features. Of the conditions with two varied features, only h0c1r1 differed from the condition with three varied features.

No significant correlations between naturalness and similarity, r_s=.01, p=.87 and preference and similarity, r_s=.05, p=.21 were found. This lack of correlations suggests that similarity did not cause the findings of Experiment 1.

In summary, people rated melody pairs with more shared features as more similar. This similarity, that can be translated to repetition within a melody, was neither correlated with naturalness of a melody, nor with preference.

	h0c0r0	h0c0r1	h0c1r0	h0c1r1	h1c0r0	h1c0r1	h1c1r0	h1c1r1
h0c0r0		.69	.82	.91	.68	.91	.90	.92
h0c0r1			.44	.66	n.s.	.71	.70	.75
h0c1r0				.37	.43	.42	.41	.52
h0c1r1					.63	n.s.	n.s.	n.s.
h1c0r0						.74	.70	.77
h1c0r1							n.s.	n.s.
h1c1r0								n.s.
h1c1r1								

Table 3: The effect sizes r (converted from the t-values) per two conditions of differences in mean similarity rating. Pairs of ratings that do not differ significantly are denoted as n.s.

General discussion

The analyses revealed that rhythm and harmony significantly influenced the naturalness rating when contour was kept constant: melodies with repeated rhythm and varied harmony were considered more natural. However, these effects disappeared when contour was varied. Only rhythm influenced the preference rating: melodies with a repeated rhythm were preferred.

As mentioned earlier, the Melody Generator II lacks rules for the continuation of the melody. The results of these experiments however, offer insight in what is needed to make a better continuation. They suggest that the rhythm of a melody should be repeated. This result is supported by claims of Brown, Merker, and Wallin (1999), who say that repetition of rhythm in music originates from the need to communicate with each other, this was presumably done with music as well as speach. Repeating a piece made it easier to recognize it and spread the music in a group of people. Moreover, repetition of rhythm can be found in many musical pieces. For example Für Elise by Beethoven, symphony 40 by Mozart and many songs in folk music. However,

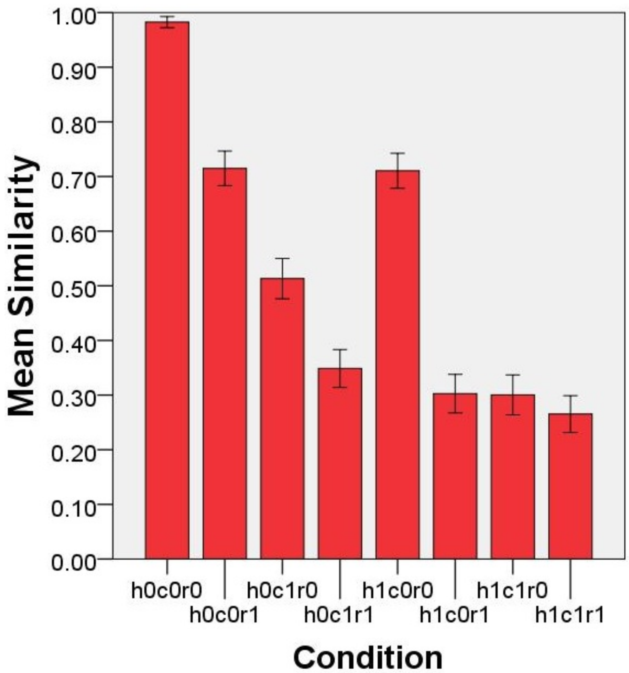


Figure 4: Average similarity of two units per condition. The first unit is a base unit, the second unit has features of the base unit varied according to the conditions as in the graph. The condition hXcYrZ is read as: harmony is X, contour is Y and rhythm is Z. The '0' means repeated, '1' is varied. The conditions are ordered according to amount of features varied. Features are ordered according to importance of results: rhythm is displayed first, then harmony then contour.

neither of the two above is direct evidence that rhythm should be repeated. The rule that follows from this research is novel: no other studies have been done about repetition, not that of rhythm, nor that of the other two features.

The results of the rhythm should however be considered with caution. In the experiment a high amount of syncopation was used and some participants pointed out that the rhythm sounded strange due to the syncopation. This might have influenced the results. On average, every condition used the same amount of syncopation per unit, but not every melody contained syncopation. The chances of a melody with a varied rhythm of receiving no syncopation were low since every unit had to receive no syncopation.

The other result concerns the harmony of a melody and translates in the following rule: when contour is repeated, the harmony should be varied. The effect of harmony on naturalness can be explained by sequence, a way of elaborating a melody. A sequence is a repeated piece of music, for example two bars, where the subsequent pieces are transposed up or down the same number of steps successively. This is similar to what is done in this experiment: the units were transposed when harmony was varied, although the number of steps and the direction was random and could differ every transition. The effect of harmony disappears when the contour is varied, which can also be explained by the sequence: if the contour is varied, there is no repeated piece to be recognized when transposing it. The effect of the sequence then disappears.

The analysis revealed that similarity was not correlated with naturalness or preference, which was in contrast with our expectations. The similarity per condition can be seen as the sum of similarity of every feature. Rhythm was rated as more natural and as preferred when more similar, while with harmony this was the other way around. So if a melody was seen as more similar the naturalness and preference ratings would go up for rhythm, but down for harmony, vice versa if the melody was less similar. The two features canceled each other out.

Future Research

Forms of repetition

The forms of repetition that were used for the experiments are by far not the only means of repeating a feature. There are more forms of repetition to be explored and possibly used by Povel's system. In this research, a structure was used where every unit was different from the last (ABCD), but there are other structures possible as well. Structures like ABAA or ABAC can emerge, where the importance lies in repetition of a part (A) that is played again after another part (B). The first part will be recognized as having already occurred, making it pleasant to listen to. This style is used very often in music and is even implemented in a way by Povel, although he only uses a small range of all possible structures. Moreover, he uses exact copies of the parts, whereas often small variations are used (e.g. in the form ABA'B, where A' is a varied form of A). Such a form of repetition is used often, for example in the nocturnes by Frédérick Chopin.

Another interesting extension of research is looking for repetition features at a much lower level. In this research the repetition units had a length of two bars, while repetitions can occur within a single bar. In the MG II it is possible to keep the rhythm the same for every bar, but even smaller repetitions can be important.

In this research only the Chord-based model was used for making the possible melodies, but as some participants made clear the use of mostly chord tones was salient and seen as abnormal. For future research it might therefore be a good idea to use the other models (Attraction-based and Scale-based) as well to test which model sounds best.

Complexity of the features

The optimal complexity hypothesis (North & Hargreaves, 1995) states that there is an inverted U-shape relation between liking of a piece of music and its subjective complexity: people like music with moderate subjective complexity best. Their results supported his hypothesis as well as results by Vitz (1966). Subjective complexity - 'how easy it is to predict what the music will do next and how many surprises the music contains' - is linked with the amount of repetition in music. The more repetition there is in music, the easier it is to predict what the music will do next. The more similar a melody is, the less complex it is. In this experiment the different combinations of the features ensured different amounts of similarity. However, no U-shaped relation (or any correlation whatsoever) was found between similarity and preference or naturalness, which would be evidence against the hypothesis.

Nevertheless, the definition of subjective complexity is focused on the music level, but it is intuitive to think there is a U-shaped relation for feature complexity and liking as well; this could explain the results. Since the rhythm was preferred when it was repeated, this could mean that the used rhythms were all relatively complex. The opposite would hold true for the harmony of a melody. For contour no relation between similarity and naturalness or preference was found, this could mean that the chosen variation was not diverse enough or that the optimal complexity of the contour was already found.

In order to find out if the optimal complexity hypothesis holds true on the feature level, these experiments should be repeated with a whole range of repetition strengths of the features.

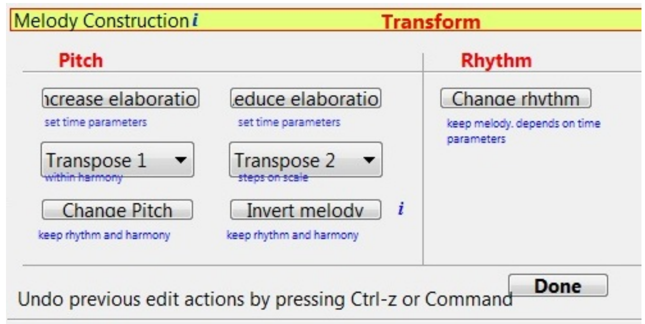


Figure 7: A Screenshot of the Melody Generator II when transforming a unit

Conclusion

Our findings suggest new rules that can be added to the knowledge based system MG II: the rhythm of a piece of music should be repeated; the harmony should be varied when the contour of the melody is repeated. Melodies that are generated at this point sound quite good although there is a strange ring to it, but with the new rules applied, the melodies will sound more natural. There are many more possible rules waiting to be found. As can be seen, the rules do not only depend on information from music theory. They can be sought in a combination of resources: music theory, but also learning from examples of existing music and extracting rules from human perception as done here. The MG II shows potential, with new rules it might compose melodies just as well as humans.

References

Brown, S., Merker, B., & Wallin, N. L. (Eds.). (1999). *The origins of music.* MIT Press.

Cope, D. (1996). *Experiments in musical intelligence.* A-R Editions.

Delgado, M., Fajardo, W., & Molina-Solana, M. (2009). Inmamusys: Intelligent multiagent music system. *Expert Systems with Applications*, *36*(3, Part 1), 4574 - 4580.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch suequences in tonal music. *Psychological Review,*, *88(6)*, 503-22.

Lerdahl, F., & Jackendoff, R. (1996). *A generative theory of tonal music.* Mit Press.

Longuet-Higgins, H., & Lee, C. (1982). The rhythmic interpretation of monophonic music. *Music Perception*, *1*, 424-441.

McIntyre, R. A. (1994). Bach in a box: The evolution of four-part baroque harmony using the genetic algorithm. *Evolutionary Computation*, *2*, 852-857.

North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*, *15*, 77-93.

Figure 5: Examples melodies for every condition. In the first condition, the first unit is repeated three times. In any other conditions a number of features have been varied. Every features that was varied, was varied three times: the second unit was different from the first, the third unit was different from the second and the fourth unit was different from the third.

Papadopoulos, G., & Wiggins, G. (1999). Ai methods for algorithmic composition: A survey, a critical view and future prospects. In *Aisb symposium on musical creativity* (pp. 110–117).

Piston, W., & Devoto, M. (1987). *Harmony.* Victor Gollancz.

Ponsford, D., Wiggins, G., & Mellish, C. (1999). Statistical learning of harmonic movement. *Journal of New Music Research*, *28-2*, 150-177.

Povel, D.-J. (2010). Melody generator: A device for algorithmic music construction. *JSEA*, 683-695.

Vitz, P. C. (1966). Affect as a function of stimulus variation. *Journal of experimental Psychology*, *71*, 74-79.

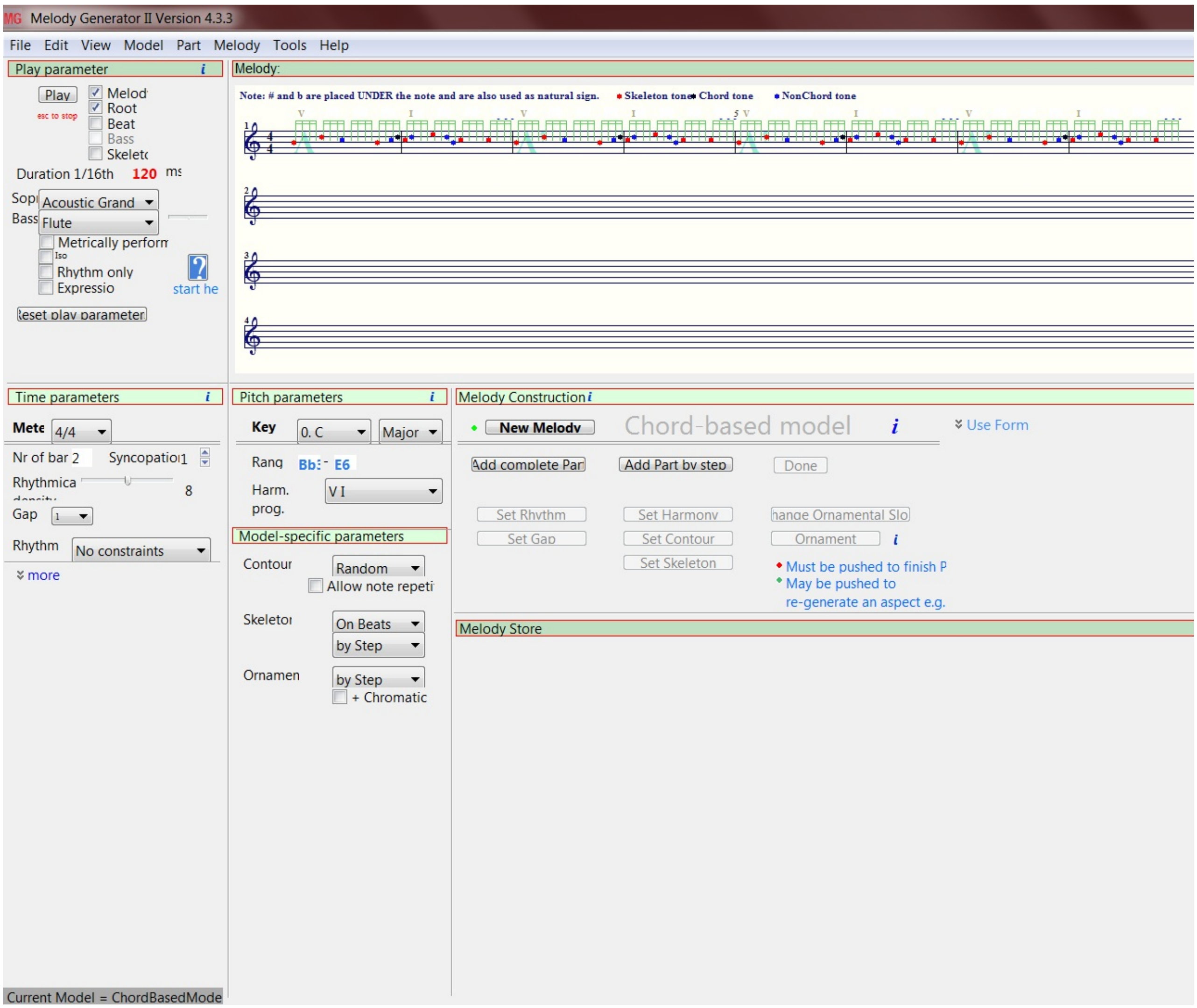


Figure 6: A Screenshot of the Melody Generator II with the basic settings that were used in this experiment

New rules for the Melody Generator II
Repeating rhythm and varying harmony
Appendix

Marieke Sweers
Dr. Makiko Sadakata
Dr. Louis G. Vuurpijl
Radboud University Nijmegen (Social science, Artificial Intelligence)

1 Materials

This section contains supplementary information about the materials used for the experiment. First, a more thorough description of the Melody Generator II and the generating of melodies will be given [1]. After that the website will be covered: the technical set-up and the format of and instruction on the website.

1.1 Melody Generator II

All melodies were made with the Melody Generator II (MGII), a program made by Dirk-Jan Povel[2]. The program was initially made to test models containing rules about how to make music to test our understanding of music. Povel reasoned that if a model could create (some basic form of) music, it would contain a valid theory of the structure of music. MG II contains rules about the structure and rhythm of music (time dimension) and about how key, harmony and tones relate to one another (pitch dimension). Povel created four different models which describe the rules of both the time dimension and the pitch dimension. These models would then be evaluated and would provide feedback to refine the algorithm and its underlying theoretical notions.

Of the four different models, the Scale-Based and Chord-Based model are the most important. The Chord-based model was explained in the main article; instead of using chord tones for the skeleton the Scale-Based model uses tones of the scale, the rest of the model is the same. As opposed to these two models, the Attraction-Based model does not first make use of the distinction between skeleton tones and ornamental tones. Instead notes are generated in sequence (from left to right) where non-chord tones attract chord tones. Lastly, the Basic

[1] The generated melodies can all be found on [makiko's website]

[2] This program can be downloaded for free from
http://www.socsci.kun.nl/ povel/Melody/index.html

model generates notes only based on the harmonic progression, using mostly chord-tones supplemented by some non-chord tones. This is the base on which the other models three build.

1.2 Generating Melodies

All stimuli (except the practice melodies) were made with the Chord-Based model. Before starting the experiment, two practice melodies were presented to make participants familiar with the stimuli. These were made with the Attraction-Based model, with a $\frac{3}{4}$ measure, to ensure participants were not biased towards the Chord-Based model. The first melody was from condition h0c0r0 (repeating all), the second from condition h1c1r1 (varying all) to make participant familiar with the whole range the stimuli could take. This was not explicitly mentioned to participants.

1.2.1 Experiment 1

Of each condition 24 melodies were generated, 2 melodies per key. Before generating melodies, several variables were set as illustrated in figure 6. Root was selected (root tone of the chord); number of bars was set to 2; gap was set to 1. Per melody one of 12 keys was chosen; syncopation, rhythmical density (chosen from 6-10) and harmonic progression were set randomly as explained in the main article. The melodies of consisted of four units. The first unit was made by clicking Add complete part and Done. The other three units were then made by duplicating the first unit: right-clicking it and selecting Duplicate Melody twice.

Depending on the condition, a different combination of features was varied. To vary a feature, the following steps were done for unit two to four. The unit was right-clicked and Transform Part was selected. A new menu now appeared bottom right (see figure 7). Contour was changed by simply clicking the Change Pitch button. Rhythm was changed by selecting a new syncoptation within the used range and a new rhythmic density and afterwards pressing Change rhythm. Harmony was varied by changing Transpose 2. The rhythm variables and the harmonic step change were chosen randomly, with the restriction for the harmonic step change that it could not be the same as the unit before. Every harmonic step change had a different change to be picked. The chance of the harmony going up or down one step was $\frac{5}{30}$, going up or down two steps $\frac{4}{30}$ continuing up to going up or down five steps ($\frac{1}{30}$ chance).

Lastly, the melody was saved by richt-clicking the melody and selecting 'Export Melody as MIDI'.

1.2.2 Experiment 2

Each trial contained two one-unit melodies: a standard melody coupled with a comparison melody. Twelve standard melodies were made, one in each key, and for each of them eight comparison melodies were made, one per condition.

The standard melody was made as the first unit in Experiment 1 and stored in the Melody store in the bottom right part of the screen by right-clicking the melody and selecting Store Melody. To make a comparison melody, the standard was selected from the Melody Store and displayed on the Melody Pane. The comparison melody was then made by varying the features according to the condition as done for the first experiment.

1.3 Web experimentation

1.3.1 Technical set-up

The website was made with html-forms with the cgi-bin construction, mainly using shellscript. In Experiment 1 the 16 melodies presented melodies were picked randomly from the whole set of stimuli, two melodies per condition not taking into account the key. Next, the melodies were ordered randomly. For experiment 2, 8 standard melodies were chosen randomly from the 12 possible with a corresponding comparison melody, where each comparison melody of 8 was from a different condition.

The response time of the participants was recorded, but not used for data processing. However, these response times might be useful for gauging whether participants were actively engaged in the experiment. If a response time was much longer than needed for choosing and selecting a rating (e.g. longer than 30 seconds), the chances are high the participant was not paying attention during the trial and its data should be removed.

There were some challenges with the website. First of all, some people were not able to play the melodies which caused a drop-out of participants. The cause was not found, but it seemed the site did work in Internet Explorer if QuickTime player was installed correctly. Nevertheless, there were some exceptions when even this did not work properly. Secondly, it was possible to go back to the previous trial and listen to a melody again, while this was unwanted. There was one participant who reported having done this.

1.3.2 The website

The website was made in two languages: English and Dutch. It consisted of three parts, experiment 1, experiment 2 and a questionnaire. At the start, participants could choose their preferred language. They were then given a consent form containing the following text:

> I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question, I am free to decline.
>
> I understand that my responses are anonymous and will be kept strictly confidential, my name and IP-address will not be collected.

> I give permission for Marieke Sweers and her supervisors to have access to my responses and agree for the data collected during this experiment from me to be used in future research.
>
> I agree to take part in the above research project.
>
> I understand the consent form and agree.

Next, they were asked for their age, native language and gender. After this, participants could check whether the melodies could be played with the MIDI sound sample and they could continu with rating practice. Experiment 1 started with an example rating form which had to be filled out and next two example trials were given which contained the two practice melodies. After this the 16 trials started. Every web page consisted of one trial containing the melody and the rating form (see figure 1). Important to note is that the two questions which were to be rated were placed below one another. It is likely this caused a bias in rating towards picking the same number for naturalness and preference.

Experiment 2 now started with two practice trials and continued with eight trials, one trial per web page (see figure 2).

In the last part, participants were asked to fill out a questionnaire about their musical background. A list of the questions can be found below

- How often do you listen to music? (hours per week)
- How often do you compose music? (hours per week)
- How often do you play musical instruments, including singing? (hours per week)
- If you play instruments, please tell us:
  1. Which instruments you play
  2. How long you have been playing / had played the instruments
  3. Your current performance level (beginner, intermediate, advanced)

2 Analysis of Comments

Of the 78 participant, 26 times commentary was given and of those 9 were positive and 4 negative. Participants reported that the experiment was fun to do and were interested in what it was about. However, people also gave some comments about the difficulty of the task and about the quality of stimulus. Some people reported that the melodies were annoying after some time.

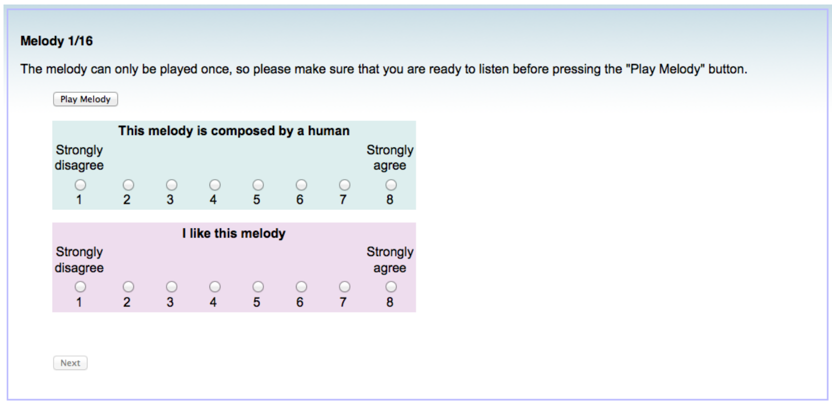


Figure 1: Web form for trials of experiment 1. Here the first trial is shown, but the form is the same for every trial.

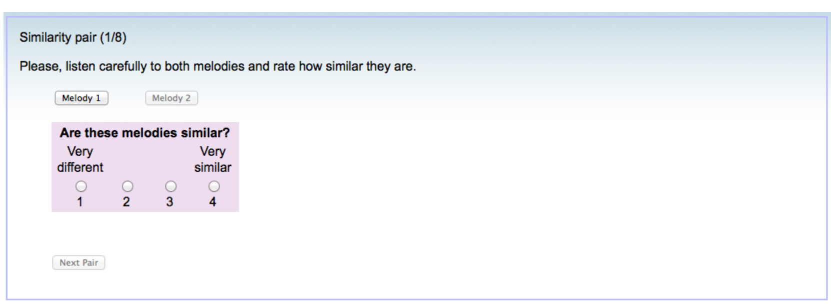


Figure 2: Web form for trials of experiment 2. Again, the first trial is shown, but the form is the same for every trial.

2.1 Task difficulty

Some participants reported that the task of determining whether or not a melody was made by a human being was difficult. They thought the task was to give their opjective judgment. However, we intended to get the subjective opinion from participants. In future research, it should be clearly mentioned a subjective opinion should be given.

2.2 Stimuli Quality

All melodies were in a MIDI format to facilitate the loading speed of the website. Most participants reported they disliked the MIDI sound. Some even reported that it interacted with their ratings: their opinion of the melody leaned more towards computer-generated than they might have if the melodies were played by natural instruments.

Participants also reported a low tone played during the melody, the root note, which was not correctly timed in the rhythm. This problem can be easily solved by selecting a different MIDI instrument for the root tone in the MG II which does not have a late onset (as the flute did have). Possibly, the root is not even needed to infer the harmony and can be left out completely, but this would first need to be tested.

3 Future research

There are many possible extensions of this research. As mentioned shortly in the article, there could be more gradations of repetition. In this research only two gradations were used: repeating or varying, while there are more possibilities. In music, many degrees of repetition are used. For example in nocturnes composed by Chopin, a theme is recognizable but it changes slightly every time it is being played. These changes are not completely changing a feature like rhythm, harmony or contour, but only changing small parts of it, thereby leaving the melody recognizable. The use of a theme troughout the music (changing a bit) is important to investigate.

When zooming in to rhythm, it is important to note that the rhythm used here was rather complicated. This could explain that participants preferred the simpler rhythm (repeated). For future research we propose using more basic rhythms. Furthermore, a single basic element (here a unit) in music does not need to remain the same length in time, instead it could be varied. For example, small variations could be made by elaborating units with ornamentations.