CORPUS-BASED RHYTHMIC PATTERN ANALYSIS OF RAGTIME SYNCOPATION

Hendrik Vincent Koops
Utrecht University
h.v.koops@uu.nl

Anja Volk
Utrecht University
a.volk@uu.nl

W. Bas de Haas
Utrecht University
w.b.dehaas@uu.nl

ABSTRACT

This paper presents a corpus-based study on rhythmic patterns in the RAG-collection of approximately 11.000 symbolically encoded ragtime pieces. While characteristic musical features that define ragtime as a genre have been debated since its inception, musicologists argue that specific syncopation patterns are most typical for this genre. Therefore, we investigate the use of syncopation patterns in the RAG-collection from its beginnings until the present time in this paper. Using computational methods, this paper provides an overview on the use of rhythmical patterns of the ragtime genre, thereby offering valuable new insights that complement musicological hypotheses about this genre. Specifically, we measure the amount of syncopation for each bar using Longuet-Higgins and Lee's model of syncopation, determine the most frequent rhythmic patterns, and discuss the role of a specific short-long-short syncopation pattern that musicologists argue is characteristic for ragtime. A comparison between the ragtime (pre-1920) and modern (post-1920) era shows that the two eras differ in syncopation pattern use. Onset density and amount of syncopation increase after 1920. Moreover, our study confirms the musicological hypothesis on the important role of the short-long-short syncopation pattern in ragtime. These findings are pivotal in developing ragtime genre-specific features.

1. INTRODUCTION

This paper presents a corpus-based study into rhythmic patterns in a ragtime corpus (RAG-collection) of approximately 11000 pieces (rags), collected by an international group of ragtime lovers [1]. The RAG-collection (RAG-C) was introduced in [16], together with an overview of open questions and a computational confirmation of musicological hypotheses of ragtime music.

Esparza et al. [3] argue that in MIR, genre classification has often been used as a proxy for measuring the success of rhythmic similarity measures, based on the assumption that *"rhythmic content is more or less homogeneous within certain musical styles"*. Their research shows that even for dance music this is often a problematic assumption. Therefore, a better understanding of the relation between rhythm and genre is important. Musicologists and ragtime fans have argued that rhythmic patterns and syncopation provide the most distinct features of the genre [1]. Edward Berlin argues that syncopation is *"at the core of the contemporary understanding of ragtime"* [2]. However, musicologists also argue that the use of rhythmic patterns has not been stable within the development of the genre over time. Therefore, we investigate ragtime's *syncopation*, its typical rhythmical *patterns* and their evolution over time.

Huron et al. [7] have shown for related genres that syncopation increases through history, something we hypothesize will be the case for ragtime as well. We reflect on the rhythmical *patterns* of the genre: what are the most characteristic rhythmic patterns used in ragtime syncopation, and does their use change over time. Berlin [2] argued for the importance of a specific short-long-short pattern in the ragtime genre, of which Volk and De Haas [16] showed that its use increased through history. We extend the research in [16] by investigating *all* patterns, to find the relative importance of this specific pattern. We hypothesize that compared to other patterns appearing in ragtime syncopation, this short-long-short pattern is one of the most characteristic patterns for the ragtime genre.

Our corpus-based study of syncopation complements extensive research on syncopation in music cognition, in which predominantly short rhythmic patterns are studied. Syncopation is considered to create violations in listeners' expectations [11], to contribute to rhythmic complexity [17] and to contribute to a sense of groove in music [12, 13]. Studying syncopation for entire compositions instead of short stimuli contributes to understanding how much violation and complexity is used in real compositions of a genre that is considered to be "highly syncopated".

Contribution. The contribution of this paper is threefold. We present a first full, systematic analysis of all rhythmic patterns in melodies appearing in a large corpus of ragtime music. Through a statistical analysis of the frequency of patterns over time, this study shows which patterns are more important in different time periods. Second, by using a formal model of syncopation, this study is able to focus on the syncopated parts of rags, commonly thought to be the most characteristic element of ragtime

[1] Ragtime Admirers Group, see http://ragtimecompendium.tripod.com/

music. Through this model, it shows the increase of syncopation use together with its most important rhythmical patterns over time. Third, a tactus finding algorithm is introduced that is capable of correctly identifying the number of beats in a bar of a ragtime piece. These three contributions are pivotal in understanding the characteristic features of ragtime music.

Synopsis. The remainder of this paper is structured as follows: Section 2 provides an introduction to ragtime music and its use of syncopation. Section 3 details the main methodology for analysing patterns and syncopation in the RAG-C. Section 4 details the results of syncopation analysis and pattern finding. The paper closes with conclusions and discussion in section 5.

2. RAGTIME

Musicologists agree that ragtime's most striking element can be found in its use of syncopated rhythmical patterns. Berlin [2] even argues that other musical features are of hardly any importance: ragtime music has no unique musical form, and its melodies do not bear any distinctive traits (except with regard to rhythm). Although rags with hardly any syncopation exist, musicologists do agree that syncopation is the dominant and distinctive element in the evolution of the ragtime genre. It is therefore that a study into ragtime will invariably involve the analysis of rhythmical patterns and syncopation.

In this research, we divide the history of ragtime music into two eras: the pre-1920 *ragtime era* and the post-1920 *modern era*. The two eras are distinguished by a remarkable increase in rhythmic experimentation and syncopation around 1920 [8, page xix]. This change was in part influenced by the French Impressionist music and piano performers mimicking the very complex rhythms of piano-roll music that were in style.

2.1 Syncopation

Syncopation is *"the displacement of the normal musical accent from a strong beat to a weak one"*, often used by composers to avoid regular rhythm by varying position of the stress on notes [14]. Musicologists have argued that ragtime's main identifying trait is its "ragged", or *syncopated* rhythm. A specific syncopated pattern is thought to be of extra importance by Harer [6] and Berlin [2]: the 'short-long-short' *121* pattern. The 121 pattern appears as ♪♩♪ in $\frac{4}{4}$ and as ♬♩♬ in $\frac{2}{4}$. The next sections detail three variants of the 121 pattern: untied, tied and augmented. Examples of these three types of syncopation in $\frac{2}{4}$ can also be found in Figure 1.

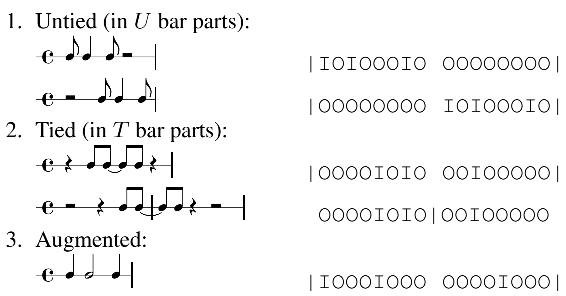

Figure 1: 121 patterns in musical notation (left) and equivalent binary onset pattern (right).

Untied syncopation. In *untied* syncopation, a pattern starts on a strong metrical position and does not pass over a bar line. In $\frac{2}{4}$, the pattern either starts on the first quarter note position or the second quarter note position. In $\frac{4}{4}$, the 121 pattern (♪♩♪) would start on either the first quarter note position or the third quarter note position. This is visualized in Figure 2 as the U bar parts. This way, the 121 pattern always constitutes the first or second half of a bar. Musicologists have argued that this type of syncopation is more characteristic of rags from the early pre-1920 ragtime era, being more prominent at the turn of the century [10], [2, p. 84].

Tied syncopation. *Tied* syncopation refers to a pattern starting on a weak metrical position. Just like untied syncopation, the tied version appears in two variants: either creating a tied note over the center of the bar, or over the barline to the next bar. This is visualized in Figure 2 as the T bar parts. In $\frac{4}{4}$ this means the pattern starts at the second or fourth quarter note position. In $\frac{2}{4}$ this means the pattern starts at the first or third eighth note position.

The *tied* pattern was found to increase during the pre-1920 era by [16]. Musicologists have argued that composers increasingly relied on tied syncopation in the late 1910s and 1920s as the ragtime style matured [10, p. 76].

Augmented syncopation. A third version of syncopation often found in ragtime music is called *augmented* syncopation. This type of syncopation augments the 121 to the length of a complete bar (3 of Figure 1). The augmented pattern appears as ♩𝅗𝅥♩ in $\frac{4}{4}$, and as ♪♩♪ in $\frac{2}{4}$. This results in a weaker syncopated pattern, which is more characteristic of early ragtime era [2, page 83], but became relatively rare after 1903.

3. METHODOLOGY

This study investigates the use of syncopation and rhythmical patterns in the RAG-C, and how these change over time. We hypothesize that syncopation is an important feature of the ragtime genre, that increases over time. To test this hypothesis, we first extract rhythmical onset patterns rags in the RAG-C, as detailed in Section 3.1. Then, to be able to group the onsets in *bars* for analysis, the number of beats per bar need to be determined. To achieve this, a tactus finding algorithm (Section 3.2) that finds the number of beats in a bar of a ragtime piece is implemented.

Differentiating between bars with and without syncopation provides insight in the patterns that are most important within ragtime syncopation. To measure the degree of syncopation of a bar, a model by Longuett-Higgins and Lee is used, as detailed in Section 3.3. These syncopation measurements are then used in a pattern recognition step (Section 3.4), to find the frequencies of all possible patterns and the relative 121 frequencies. The following sections describe each of these steps in detail.

3.1 Onset Extraction

Characteristic of ragtime music is a ragged or syncopated melody over a stable accompaniment that reinforces the meter. The importance of first separating a piece into its individual rhythmic layers for syncopation measurements was shown in [13]. Therefore, to be able to analyse syncopation of the melody of a rag, we split it from its accompaniment. The accompaniment is used in a tactus finding step (detailed in Section 3.2), and the melody is used in a pattern finding step (section 3.3).

The melody and accompaniment are split using the skyline with dip-detection method detailed in [15, 16], which performs a near-perfect splitting of a melody and its accompaniment on a subset of the RAG-C. To be able to analyse rhythmical patterns properly, both the melody and accompaniment are quantized. We use the technique described by Volk & De Haas [16], with the exception of using four bins per quarter note, instead of twelve. This results in quantisation to a sixteenth note grid, which we can apply to the formal model of syncopation described in Section 3.3. Because of different *normalized average quantisation deviations* (the average deviation of notes divided by the MIDI quarter note length, see [16]) between files in the dataset, we keep track of the quantisation error, and disregard all MIDI files with a normalized average quantisation error above 2%.

This results in two sequences of onsets per rag, one representing the rhythm of the melody, and one representing the rhythm of the accompaniment. The onsets in the sequences are represented with `I`'s as sounding events and `O`'s as non-sounding events. See the bottom two rows of the tree in Figure 2 for an example with its music notation equivalent.

3.2 Tactus Finding

This study analyses the onset patterns that appear in syncopated *bars* of rags. The method in Section 3.1 results in a sequence of onsets, therefore we need a way to segment this sequence into bars. One way to achieve this is to use the annotated MIDI time signature of the rags, but from a manual inspection this information was found to be not always reliable. Therefore, a tactus finding algorithm is created that is able to find the number of beats in a bar. This information is used to group the right number of onsets into bar representations: from a sequence of onsets to segments representing bars.

Two features of ragtime music facilitate time signature detection from onset patterns with greater ease, compared to other genres. First, most rags are written in either $\frac{2}{4}$ or $\frac{4}{4}$, other meters are rare. Secondly, a characteristic feature of ragtime is a stable metre pattern in the accompaniment underneath a syncopated melody [9]. As a general rule, the accompaniment *"reinforces the meter with a regular alternation of low bass notes or octaves on the beat, alternating with mid-range chords between the beats"* [1].

In $\frac{4}{4}$ (and $\frac{2}{2}$), this alternation appears as ♩♩♩♩. In $\frac{2}{4}$, this pattern appears as ♫♫. This pattern can be used to estimate the number of beats in a bar for duple time signatures.

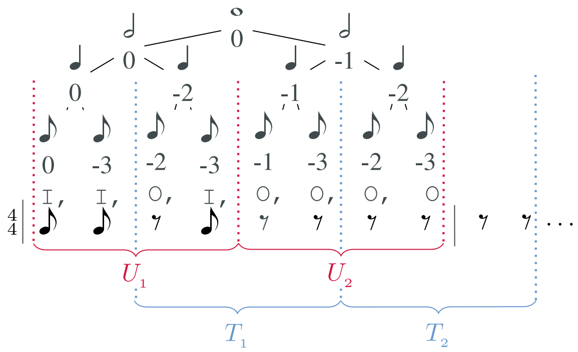

Figure 2: Example of a hierarchical metric tree with values. U and T denote the bar parts where we search for patterns. The onset pattern (`I`'s and `O`'s) represents the 121-pattern (♪♩♪) in a U part of a bar. The LHL value of this bar is $((-2) - (-3)) + ((-1) - (-3)) = 1 + 2 = 3$.

The algorithm presented in this paper finds the number of beats of a rag by assuming that in $\frac{2}{4}$ the onset density in the accompaniment is higher than in $\frac{4}{4}$. The onset density d for a sequence of onsets is calculated by dividing the number of onsets on even positions by the number of onsets on odd positions. If d is larger than a certain threshold, it is assumed that the onset density is low, and a time signature of $\frac{4}{4}$ is assumed. If the fraction is lower than a threshold, the onset density is high and the time signature is assumed to be $\frac{2}{4}$. Using this information, the onset sequence is either segmented in 16 onsets per bar in the case of $\frac{4}{4}$ and 8 onsets per bar in the case of $\frac{2}{4}$. These bar onset patterns are then used in a next step in which the amount of syncopation is measured, as explained in Section 3.3.

Evaluation. The tactus finding algorithm is evaluated in an experiment using 200 randomly selected rags from the RAG-C. After quantization and selecting rags with a normalized average quantization error below 2%, 72 rags remain. The rags are manually annotated with their correct time signature by a music expert. Using the technique described in 3.2, the algorithm predicts the correct number of beats in a bar in 92% (66) of the rags using a threshold of $d = 0.8$. Of the six songs that were incorrectly identified, two are not in a $\frac{2}{4}$ or $\frac{4}{4}$ time signature ($\frac{3}{4}$ and $\frac{6}{8}$) and four lack the typical accompaniment pattern. These results show that this method is highly useful as a preprocessing step for segmenting onsets into bars.

3.3 Longuet-Higgins & Lee Syncopation Measurement

For pattern analysis, we differentiate between bars with and without syncopation and analyse the former, to find its most characteristic patterns. A formal model of syncopation introduced by Longuet-Higgins & Lee (LHL) [11] provides a numerical representation of syncopation in a bar by assuming that a rhythm in a meter is interpreted by a listener by minimizing the amount of syncopation. In an experimental comparison between different syncopation measurements, Goméz et al. [5] found that the LHL agrees closely with the human judgement of syncopation. The notion of minimizing syncopation is expressed in the algorithm, in which syncopation is defined to occur when a note occurs on a weaker position than its succeeding rest (or tied note). This was also shown empirically by Fitch et al. [4], who showed that the recall of a rhythm decreased with higher LHL syncopation.

The LHL model computes syncopation using a tree of metric hierarchy (see Figure 2 for an example). This tree is built to a minimal depth needed to represent the notes. For example, if a $\frac{4}{4}$ bar only contains two half notes (𝅗𝅥𝅗𝅥), a tree of depth 1 is used. In case a note appears on a deeper level, a deeper tree is used (e.g. depth 4 in 𝅗𝅥𝄾𝅘𝅥𝅯𝅘𝅥).

The nodes of the tree are populated with values k given to the left children and $-d$ to right children, where k is the value of the parent of a node and $-d$ is the negative value of the depth of the tree at that node. The value of the root of the tree is 0.

In the LHL model, syncopation occurs where a note (I) with a lower value is followed by a rest (O) with a higher value. The example in Figure 2 contains two of these (I, O) pairs, the second eighth note followed by a rest, and the third eighth note followed by a rest. The amount of syncopation for a pair is the difference in values: O-I, $(-2)-(-3) = 1$ for the first example. The total syncopation value of an entire bar is the sum of syncopation pairs within that bar:

$$\sum_{i=1}^{n}(\nu(\mathtt{O}_{\mathtt{i+1}}) - \nu(\mathtt{I}_{\mathtt{i}})) \quad \text{if} \quad \nu(\mathtt{I}_{\mathtt{i}}) < \nu(\mathtt{O}_{\mathtt{i+1}}) \qquad (1)$$

where the subscript denotes the i^{th} position in the bar of length n and $\nu(\varphi)$ denotes the metric tree value of φ.

3.4 Pattern finding

To find the frequencies of onset patterns in the RAG-C, a pattern finding algorithm is created. We are interested in the bar parts where the tied, untied and augmented 121 pattern can appear. Therefore, this algorithm finds the frequency of candidate patterns in U and T bar parts (see Figure 2). With this quantitative measurement of pattern frequencies, we measure whether the 121 pattern is indeed characteristic for ragtime music in these bar parts, and what other patterns are important. To be able to search for patterns in U bar parts, each bar from the RAG-C is concatenated with half of the bar that follows it.

To find the frequencies of patterns in U and T, all possible combination of I's and O's' are generated for the length of half a bar. For example, in the case of a bar in $\frac{4}{4}$ quantized on sixteenth notes, a full bar contains 16 onsets. Therefore, all candidate patterns (Π) of length 8 are generated: [O,O,O,O,O,O,O,O]...[I,I,I,I,I,I,I,I]. The frequency of each candidate pattern $\rho \in \Pi$ is calculated by counting how often each one appears in one of the U and T parts, normalized over the total number of bars. Calculating the frequency of all pattern results in distributions of patterns in U and T bar parts.

4. RESULTS

This section describes statistics of syncopation and the results of finding the most frequently used patterns in U and T parts of syncopated bars. First, results of finding the most frequent patterns in the entire RAG-C are presented (Section 4.1). Then, the RAG-C is split into rags from the ragtime era (before 1920) and the modern era (after 1920) to find which patterns are characteristic for these eras. These results are presented in Section 4.2. In the next sections, $\bar{x}$ denotes an average and σ denotes a variance.

4.1 Syncopation in the RAG-C

From the RAG-C, 356519 bars are extracted, of which 46% (163197) are syncopated (i.e. LHL > 0). The average LHL value of syncopated bars is $\bar{x} = 2.02$, $\sigma = 1.08$. The largest LHL syncopation is 15, corresponding with only 28 bars in the RAG-C. Nevertheless, a little over half of the bars (54% = 193322) in the RAG-C is devoid of any syncopation (i.e. LHL=0).

Finding the most frequenly used patterns in T and U bar parts of bars with LHL$>$ 0 yields the results in Figure 3. Note that patterns are part of a syncopated *bar*, and not necessarily syncopated themselves. For example, a bar consisting of |IOIOOOIO OOOOOOOO| is syncopated because of the 121 pattern in the first half of the bar (IOIOOOIO), however, the second half (OOOOOOOO) is devoid of any syncopation.

The figure shows that the 121 pattern appears as the most frequent pattern in T bar parts, and as a third most frequent pattern in U bar parts. This affirms the hypothesis that when taking rags from all time periods in consideration, the 121 pattern is indeed one of the most characteristic ragtime patterns. The figures show that for the whole RAG-C, the 121 is more characteristic in T than in U.

Ragtime and Modern era. We split the RAG-C into pre-1920 *ragtime era* bars and post-1920 *modern era* bars,

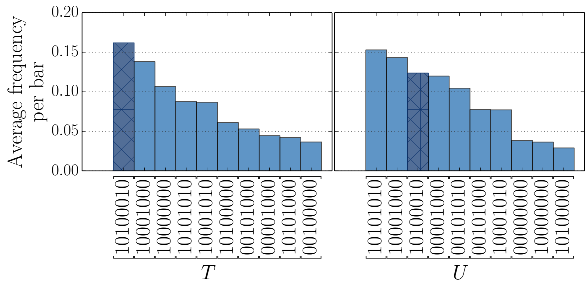

Figure 3: The 10 most frequent patterns in T and U parts of bars with LHL > 0. The 121 pattern is visualized darker.

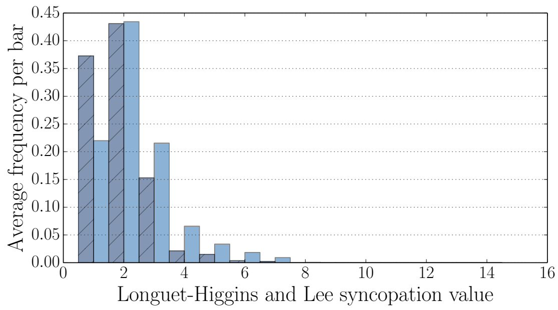

Figure 4: Percentage of bars with LHL$>$0 in ragtime era (dark) and modern era (light). Values $>$ 7 are too small to be visible.

to find the change in syncopation degree over time. The average LHL syncopation of a ragtime era bar with LHL>0 is $\bar{x} = 1.9$, and $\bar{x} = 2.4$ in the modern era. In a Wilcoxon test for the null hypothesis that two related paired samples come from the same distribution, we find that $p \ll 0.001$, which shows that the modern era is significantly stronger syncopated. Also taking into account the non-syncopated bars shows a significant difference, with $\bar{x} = 0.83$ (ragtime era) and $\bar{x} = 1.26$ (modern era), again with $p \ll 0.001$.

Figure 4 shows distibution of LHL syncopation found in syncopated bars from these two eras. The figure shows that bars with LHL=1 are more common in the ragtime era, and LHL=2 is almost equally common in the ragtime era as in the modern era. Nevertheless, it also shows that bars with LHL>3 are more characteristic for bars from the modern era. Bars with LHL>5 occur twice as often in the modern era compared to the ragtime era.

Syncopation per rag. To find the distribution and degree of syncopated bars of complete rags in the RAG-C, we computed statistics on rags. The average syncopation per rag for the whole RAG-C is $\bar{x} = 0.95$, $\sigma = 0.6$. An LHL value of 1 roughly corresponds with a single syncopation inside one of the U parts, resulting in a bar of ♩♫𝄽–.

For the ragtime era, the average syncopation per rag is $\bar{x} = 0.85$, $\sigma = 0.52$. For the modern era this is $\bar{x} = 1.28$, $\sigma = 0.74$. Therefore, in the modern era, syncopation more often appears on weaker metric positions that correspond with lower values in the LHL tree, thereby increasing the LHL value of the bar. An example of this is 𝄾♩𝄽♩𝄾. For both eras, we find that the number of syncopated bars per rag is around 50%, which means that not the number of syncopated bars increases, but the use of syncopation inside bars does. We found that the difference in syncopation between ragtime and modern era to be highly significant with $p \ll 0.001$. When only taking into account the syncopated bars, we find $\bar{x} = 1.84$, $\sigma = 0.54$ per rag for the ragtime era, and $\bar{x} = 2.29$, $\sigma = 0.67$ per rag for the modern era, again with $p \ll 0.001$.

Both the statistics on rags and bars show that overall, stronger syncopation is more characteristic of modern era rags. In the modern era, syncopation occurred more often on weaker metric positions, thereby increasing the LHL syncopation. The next section details the difference in patterns found between these eras.

4.2 Frequent Patterns in Ragtime and Modern Era

To find a change in pattern use in syncopation over time, we look at the patterns found in syncopated bars from the ragtime era and modern era. Figure 5 and Figure 6 show the 10 most frequent patterns found in U and T bar parts In the figures, the 121 pattern is visualized darker.

Patterns appearing in U bar parts. The left side of Figure 5 shows that the 121 pattern in U occurs more frequently in the modern era compared to the ragtime era. Secondly, it shows that the 121 pattern in U also became more important over time compared to other patterns. Although the 121 pattern in the modern era is the second most frequent pattern, the difference between the first and third

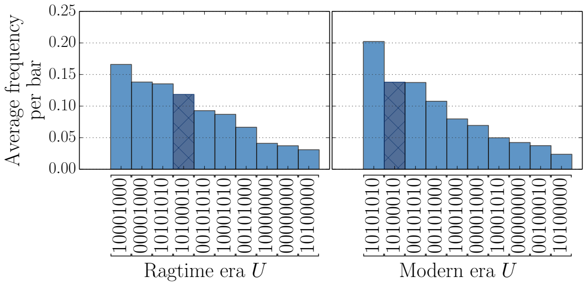

Figure 5: The 10 most frequent U patterns found in bars with LHL > 0 in ragtime and modern era. 121 pattern is visualized darker.

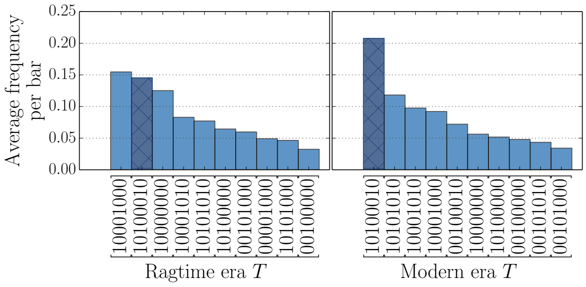

Figure 6: The 10 most frequent T patterns found in bars with LHL > 0 in ragtime and modern era. 121 pattern is visualized darker.

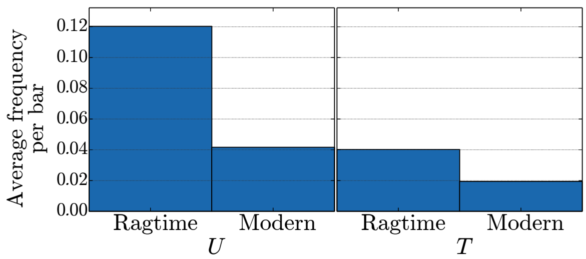

Figure 7: Frequency of augmented pattern in U (left) and T (right) bar parts of ragtime era and modern era bars.

most frequent pattern (𝄾♪♫) is minimal, The figure affirms the hypothesis that the 121 pattern is an important pattern amongst other patterns, and its importance for the ragtime genre in U increased over time.

Another difference between the ragtime and modern era can be found in the onset density and metrical position of onsets. In the ragtime era, the top most frequent patterns have a lower onset density and have more notes on strong metrical position, indicating that the patterns used in the ragtime era are less complex. In the modern era, the top most frequent patterns are more dense and have more notes occurring on weaker metrical positions.

Patterns appearing in T bar parts. Figure 6 shows the 10 most frequent patterns found in T parts of syncopated bars. The figure shows that the 121 pattern is an important pattern in T, being the second most frequent pattern in the ragtime era and by far the most frequent pattern in the modern era. These results affirm the hypothesis that both the importance and use of the 121 pattern in T has increased over time. In a study by Huron et al. [7], a top 10 of most frequently found syncopation patterns in sound recordings of American popular music spanning the period

1890 to 1939 is presented. The most frequently found syncopation pattern appears here as well, as the fifth most important pattern in the modern era bars: 𝄾♩♩♪ (OOIOOOIO).

Furthermore, an increase in onsets on weak metrical positions is observed. The first couple of most frequent patterns in the ragtime era are simple rhythms on strong metrical positions. Conversely, in the ragtime era we observe denser patterns. Both the increase of 121 use and use of denser patterns is in line with the argument of Jasen [8] that after around 1920, rags became more difficult to play, because *"[...] writers were no longer writing for the at-home amateur pianist, [...], but were writing for themselves and for other professional performers"*.

U and T patterns between eras. A comparison between the leftmost figures of Figure 5 and Figure 6 shows that for both T an U, the most frequent ragtime era pattern is the same regular sparse pattern (𝅗𝅥𝅗𝅥). Since this pattern itself is not syncopated according to the LHL model, this shows that in syncopated bars from the ragtime era most often only one half of a bar contains syncopation, indicating a lower amount of overall syncopation compared to the modern era. The 121 pattern is an important pattern for the ragtime era, being the fourth most important pattern in U bar parts and second most important pattern in T bar parts. The use of the 121 pattern increased over time, both in U as in T. In the modern era, the 121 pattern is by far the most important pattern in T bar parts, and the second in U.

Overall, the most frequent patterns in the ragtime era show more sparse onset patterns on strong metrical positions, indicating simpler rhythms. The most frequently observed pattern in the ragtime era (♩♩) corresponds with a part of the third most common syncopation pattern found by Huron et al. [7] in sampled music from 1890 to 1939. Nevertheless, the research by Huron et al. does not focus specifically on ragtime music, so further cross-genre research is needed to find if this pattern is specifically important for ragtime. Conversely, it is observed that the patterns in the modern era are more dense. Onsets appear more frequently on weaker metrical positions, increasing the complexity of patterns in terms of onset density over time. This agrees with the musicological hypotheses that earlier ragtime is simpler, and the exceptional renewed rhythmical creativity from around 1920 onwards [2, 8].

Augmented syncopation. Figure 7 shows the frequencies of the augmented 121 pattern in U and T bar parts. In U, a difference of around 60% is observed, which reflects the argument by Berlin [2] that the pattern becomes "quite rare" at the end of the ragtime era. Although rare to begin with in T, the occurrence drops with 50% in the modern era compared with the ragtime era. Care should be taken with drawing conclusions from these results because of the low frequency. The observations on the augmented pattern underline the overall trend of ragtime moving towards using onsets on weaker metrical positions and increased onset density of patterns, thereby becoming more syncopated.

5. DISCUSSION AND CONCLUSION

Through this study, we were able to confirm new and existing hypotheses on increasing syncopation and rhythmic pattern use in the ragtime genre.

Ragtime music is often described as 'highly syncopated'. Through the RAG-C, we showed for the first time that in a large corpus this translates into about half of the bars of rags being syncopated. Musicologists have argued that syncopation is important for the ragtime genre. Through the computational means in this study, we can affirm the hypothesis syncopation is a characteristic feature of the genre. We can also confirm the hypothesis that the amount of syncopation is not stable over time, but increased after 1920. More specifically, by exploring this notion of increased syncopation we discovered that the number of syncopated bars is approximately equal in ragtime and modern era rags, but that the LHL values of bars increases.

In an analysis of all patterns used in ragtime syncopation, we showed the top 10 most frequently used patterns in syncopated bars. We found that over time, onset patterns became more dense with more notes on weaker metrical positions. This finding is consistent with the increase of LHL. We can affirm the findings by Volk and De Haas [16] on the increase of 121 after the ragtime era. In addition, we showed that the 121 pattern is a highly important rhythmical pattern for the genre, being one of the most frequently used patterns compared to all other patterns.

Our corpus-based study on syncopation complements studies in music cognition research, which have investigated syncopation's role on violating listeners' expectations, thereby contributing to listening pleasure of the music [17]. These studies are predominantly carried out on short rhythmic stimuli. Understanding the full power of syncopation requires its study within entire compositions as realized within this paper. Violating listeners' expectation through the use of syncopation in this ragtime corpus is realized on average in half of the bars in the melody. Whether or not there are other genres that use even more violations, while still providing a clear sense of meter, will have to be addressed in future research.

To study what 'highly syncopated' means in the context of other genres, we plan on comparing the amount of syncopation in the RAG-C to other genre datasets. Furthermore, a study into the use of the 121 pattern in other genres would shed light on the relative importance of the pattern to ragtime and other genres.

6. ACKNOWLEDGEMENTS

We thank R.C. Veltkamp and D. Bountouridis for providing valuable comments on an earlier draft on this text. The authors would like to thank anonymous reviewers for their valuable comments and suggestions. H.V. Koops, A. Volk and W.B. de Haas are supported by the Netherlands Organization for Scientific Research, through the NWO-VIDI-grant 276-35-001 to A. Volk.

7. REFERENCES

[1] Edward A. Berlin. *Ragtime*. Oxford Music Online. Grove Music Online. Oxford University Press. Addressed: April 24, 2015.

[2] Edward A. Berlin. *Ragtime: A musical and cultural history*. Univ of California Press, 1984.

[3] Tlacael Miguel Esparza, Juan Pablo Bello, and Eric J. Humphrey. From genre classification to rhythm similarity: Computational and musicological insights. *Journal of New Music Research*, 44(1):39–57, 2015.

[4] W. Tecumseh Fitch and Andrew J. Rosenfeld. Perception and production of syncopated rhythms. *Music Perception: An Interdisciplinary Journal*, 25(1):pp. 43–58, 2007.

[5] Francisco Gómez, Eric Thul, and Godfried T Toussaint. An experimental comparison of formal measures of rhythmic syncopation. In *Proceedings of the International Computer Music Conference*, pages 101–104, 2007.

[6] Ingeborg Harer. *Ragtime: Versuch einer Typologie*. Schneider, 1989.

[7] David Huron and Ann Ommen. An empirical study of syncopation in american popular music, 1890–1939. *Music Theory Spectrum*, 28(2):211–231, 2006.

[8] David A. Jasen. *Ragtime: an encyclopedia, discography, and sheetography*. Taylor & Francis, 2007.

[9] Samuel A. Floyd Jr. and Marsha J. Reisser. The sources and resources of classic ragtime music. *Black Music Research Journal*, 4:pp. 22–59, 1984.

[10] Stanley V. Kleppinger. On the influence of jazz rhythm in the music of Aaron Copland. *American Music*, 21(1):pp. 74–111, 2003.

[11] H. Christopher Longuet-Higgins and Christopher S. Lee. The perception of musical rhythms. *Perception*, 11(2):115–128, 1982.

[12] Guy Madison and George Sioros. What musicians do to induce the sensation of groove in simple and complex melodies, and how listeners perceive it. *Frontiers in Psychology*, 5(894), 2014.

[13] George Sioros, André Holzapfel, and Carlos Guedes. On measuring syncopation to drive an interactive music system. In *Proceedings of the 13th International Society for Music Information Retrieval Conference, ISMIR 2012, Porto, Portugal, October 8-12*, pages 283–288, 2012.

[14] Syncopation. *The Oxford Dictionary of Music, 2nd ed. rev*. Oxford University Press, 2006.

[15] Alexandra L. Uitdenbogerd and Justin Zobel. Manipulation of music for melody matching. In *Proceedings of the sixth ACM international conference on Multimedia*, pages 235–240. ACM, 1998.

[16] Anja Volk and W. Bas de Haas. A corpus-based study on ragtime syncopation. In *Proceedings of the 14th International Society for Music Information Retrieval Conference, ISMIR 2013, Curitiba, Brazil, November 4-8, 2013*, pages 163–168, 2013.

[17] Maria A.G. Witek, Eric F. Clarke, Mikkel Wallentin, Morten L. Kringelbach, and Peter Vuust. Syncopation, body-movement and pleasure in groove music. *PloS one*, 9(4):e94446, 2014.