HOW DIFFERENT ARE RAP FLOWS? COMPARATIVE ANALYSIS ON LANGUAGE AND GENRE.

Haeeun Kim haeeun.kim@epfl.ch
Frank Dupont frank.dupont@epfl.ch
Graziano Conti Rossini graziano.contirossini@epfl.ch

ABSTRACT

In vocal artistry, lyrics serve not only the content of the song. The intrinsic sonic elements of lyrics, such as pronunciation, is delivered through lyrics. Being a part contemporary music with lyrics and also a part of Hip-Hop culture, rap is the association of music and poetry elements and is greatly influenced by street culture. Culture being different depending on time and location, this work aims to compare various sonic components of rap flow in Hip-Hop music, between different languages and genres. Using a humdrum-formatted corpus of 52 American rap songs and 10 French rap songs, we conduct comparative analysis on the tempo, rhyme, and rhythm of flow in American Oldschool, American Newschool, and French Oldschool raps. Results reveal that the musical characteristics of rap flows differ more greatly between languages than between genres.

1. INTRODUCTION

Rap is a musical form of vocal delivery, a vocal expression which is usually performed over a backing track or an instrumental. It can be traced back to its African roots centuries before the contemporary period [8]. Rap first appeared in the United States in the 1970s to then spread around the world [8], also appearing in some Jazz or Blues songs [1]. Nowadays, rap is commonly associated with Hip-Hop music, which we specifically cover in this study. Rap has three main components which are the “content” (the lyrical text itself), the “flow” (the use of rhythms and rhymes), and the “delivery” (the vocal techniques to perform the rhythms and rhymes) [5].

Although these three components are strongly linked, our study focuses on the “flow” component because we think it is the most significant characteristic of rap music compared to other musical genres and it is a good starting point for describing rap music. The term "flow" in rap music is not formally defined, but there is a consensus that the rhythm and the rhymes are important for describing flow [5].

In recent works about flow, Mitchell Ohriner investigates general beat characteristics of the genre focusing more on a specific track and the difference in signing by featuring MCs, based on OCR technique" [7]. Kyle Adams also defines the metrical techniques used by different individual rappers from different styles [2]. But one of the most complete study is made by Nathaniel Condit-Schutz from the University of Ohio State called "MCFlow: A Digital Corpus of Rap Transcriptions" [3] which mainly covers some general descriptors of “flow” and the evolution of these descriptors over time. We build on these works by using their features to look at differences between genres and across languages.

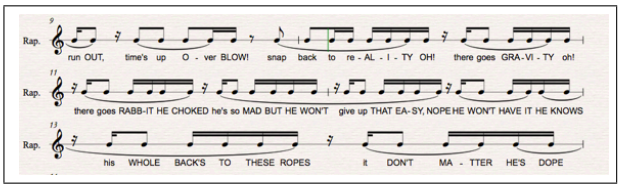


Figure 1: Rap flow diagram of *Lose yourself*, by Eminem. Representation in Western musical notation.

2. KEY ELEMENTS OF FLOW

Since the flow concept has multiple definitions, we will introduce the definition of rap flow we use for our research: Flow are musical experiences expressed in rhyme, rhythm and stress, delivered in sentences that are in accordance to the instrumentals [5]. Figure 1 shows an example of a rap flow diagram represented in Western musical notation [4]. It is a score where the temporal position of each syllable is given using [Western] rhythmic notation together with its [the syllable’s] text. Slurs between a series of notes correspond to the start and end of each phrase and the capitalized syllables indicate an occurring rhyme.

From this flow diagram, we can extract the following key elements of flow.

- Tempo: The speed in which instrumentals and lyrics are delivered is an important musical element that affects an individual’s emotional experience [11].
- Rhyme: Being an essential element in Western lyrical music and poetry, the importance of rhyme is explicitly emphasized in rap practice.
- Phrasal boundary: Phrases and sentences in lyrics work as semantic units of songs. Since phrasal boundaries usually coincide with marked rests, emcees ’breathe’ at the phrasal boundaries.

- Rhythm: Being an essential element in any musical style, the importance of rhythm is emphasized in rap practice, where pitch structure is de-emphasized.

Although its concept of repetition of similar pronunciations is shared across languages, rhyme reflects cross-linguistic differences [6]. Thus we can expect to observe different sonic experience in songs from different languages. On the other hand, it is a common knowledge that in any musical style, there are similarities and differences between detailed subdivision of genres. Since Hip-Hop is a musical practice that is found in different languages and has a lot of subdivision of genre, we are motivated to associate sonic characteristics of flow to its linguistic characteristics and the particularities between genre. In this study, we narrow our focus to only two languages, which are American and French, and two genres, which are "Old-school" and "Newschool" (or "Gangasta.)

We make three hypotheses. (1)The flow of the Old-school is faster than the Newschool. Instrumentals and lyrics are delivered at a faster tempo in French and American Oldschool than in American Newschool. (2)The flow of the American rap uses rhyme more often and diversely, compared to the flow of French rap. (3) The American Oldschool flow rhythm is more similar to that of French Oldschool flow rhythm than that of American Oldschool flow rhythm. Inter-genre difference is greater than inter-ligual difference.

3. DATASET

Our dataset is composed of two parts: (1) Two sets of 26 transcriptions of American raps taken from the MCFlow corpus [3]. (2) A set of 10 newly created transcriptions of French raps.

In order to compare flow across different genres, we classify each song in the dataset to their corresponding genres. The classification is based on the year of release of each song and our expert judgement. We make the assumption that the Oldschool and Newschool genres appeared in different time periods. Oldschool appears during the 90's while Newschool appears later around the beginning of the 21st century. We also verify this initial classification with the classification given by the English Wikipedia entries. Only the songs with a 4/4 time signature are selected.

Finally, we have 62 songs in three groups: 26 songs for the American Oldschool genre, 26 songs of the American Newschool genre and 10 songs for the French Old-School genre.

3.1 MCflow Dataset

The MCFlow dataset contains 124 famous American rap songs, selected by their popularity in the Top 100 chart, in the Humdrum format. We use the following flow descriptors according to the definition in [3], omitting some of the descriptors that were not relevant for the present study.

- Surface descriptor: Marks the temporal (and metric) position of notes and their associated duration in Western musical notation.

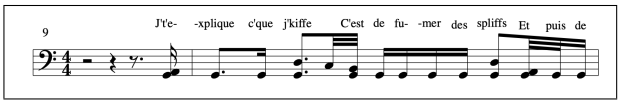


Figure 2: An excerpt of manually encoded version of *That's my people*, by NTM

- Break descriptor: Defines the boundaries between prosodic units and marks the separation between two phrases.
- Rhyme descriptor: Marks if each note is associated with rhymes. In this dataset, assonance, alliteration and slant rhymes are noted as well [3].
- Lyrics descriptor: Marks associated lyrics for each note. This provides some basic information about syntactic boundaries as well.

3.2 French song transcription

We select 10 songs by their popularity in the French Top 50 chart [9] and by the 10 most sold rap albums in francophone countries [10]. This dataset is for now small compared to the American one but we encourage future researches to add new transcriptions to the current one.

There are several steps in our encoding which leads to the final French transcription. For each syllable, we use a rhythmic notation to express the temporal information and pitch to encode different features. We assume that the tempo doesn't change within the song. The transcription is stored in MusicXML format, which can be converted to a representation that is similar to the Humdrum format used in [3]. An example of MusicXML transcription represented in Western score is shown in Figure 2.

4. METHOD

4.1 Data Preprocessing

We convert entries of humdrum-formatted data into formats that can be directly computed. We translate beat lengths written in musical notation into relative beat length. If a note written in '16', which indicates a sixteenth note beat, will be translated into '0.25', since all the songs in the dataset have 4/4 meter. Rest notes are merged with the preceding syllables. Finally, We add "position" column that adds up the syllable lengths("**length") up to each row, starting from the beginning of the song. As a result, for each syllable, we have length and position represented relative to reference beat and rhyme and phrasal boundary indicators. An example is shown in Table 1.

4.2 Formal definition of flow features

We formally define flow features based on the selected descriptors in addition to metrical tempo. For each feature, we consider its intrinsic properties and apply different levels of locality. In this study, we use phrases as the smallest unit of locality, which is separated by 'break' descriptors.

- Phrase level: Flow characteristic of each phrase is independent to those of neighboring phrases.

**surface	**lyrics	**length	**position
16	J't'e-	0.25	4
8.	-xplique	0.75	4.75
16	c'que	0.25	5
8..	j'kiffe	0.875	5.875
32	C'est	0.125	6
16	de	0.25	6.25
16	fu-	0.25	6.5
16	-mer	0.25	6.75
16	des	0.25	7
8	spliffs	0.5	7.5
32	Et	0.125	7.625
32	puis	0.125	7.75
16	de	0.25	8

Table 1: An excerpt of *That's my people*, by NTM, up to the second step of pre-processing. Rhyme and break indicators are not shown in this table.

- Song level: Flow characteristic of each phrase is either identical or related to those of neighboring phrases of the same song, but not by other songs in the same group.
- Group level: Flow characteristic is a general description of the selected feature in each group.

4.2.1 Tempo

In this study, we consider metrical tempo and syllabic tempo (or syllabic density). Since rap is a lyrical musical genre, it is important to note syllabic tempo separately from the metrical tempo. While metrical tempo does not change within a song in contemporary popular music, syllabic tempo may vary. A song may contain both the "talkative" part and "taciturn" part at the same time. For metrical tempo, we follow song-level definition. For syllabic tempo, we keep the variability between phrases, by defining the tempo on the phrase-level, whose formal representation is shown below:

$$SyllabicDensity = \frac{T}{60} \cdot \frac{n}{B}$$

where T is the metrical tempo of the song, n is the number of syllables in a phrase, and B is the phrase length represented in beats.

4.2.2 Rhyme

Since we are focusing on sonic elements of rhyme, we start our analysis on rhyme by representing different rhyme practices by their frequency(or density) and variability of positions. Although rhyme is a local phrasal property, it occurs in relation to neighboring phrases. Considering this intrinsic relational property of rhyme, we apply song-level definition to rhyme density. In addition, rhyme is usually aligned with the underlying meter of each song, which may vary temporally according to its metrical tempo. In this study, by choosing measure as the unit of rhymes, we focus on metrical property than on temporal property of rhyme. The formal representation of rhyme density is shown below:

$$RhymeDensity = \frac{N}{M}$$

where N is the number of total occurrences of rhyme in a song, and M is the number of measures in the song. For the variability of rhyme positions, we apply the entropy-based analysis done in [3], where variability of rhyme position is defined as entropy of metric positions of rhymed syllables. Since our question is how rhyme practices vary between different groups of songs, not between different songs and emcees within a selected group, the operation is applied to all the rhymed syllables in each group. Metrical position of rhymed syllables are obtained through applying modulo the length of the bar to the ***position* descriptor. The formal definition is shown below:

$$RhymeVariability = H(P)$$

where P is the distribution of metrical positions of rhymed syllables in each group.

4.2.3 Rhythm

So far, we defined the features that give a general representation of flow, assuming the full independence of each feature. In rhythm analysis, we model sequential structure of syllable duration within phrases for each group, using bigram transition probability and 1st-order Markov model. In order to represent how the rhythm of phrases starts and ends, we insert phrasal boundary indicators, 'start' and 'end', at each phrasal boundary. The first syllable of each phrase always follows the 'start' indicator and the last syllable of each phrase is always followed by the 'end' indicator. We aggregate all possible bigrams and count their occurrence in each group. For a bigram (*i*,*j*), where *i* and *j* are syllable duration, the transition probability is defined as following:

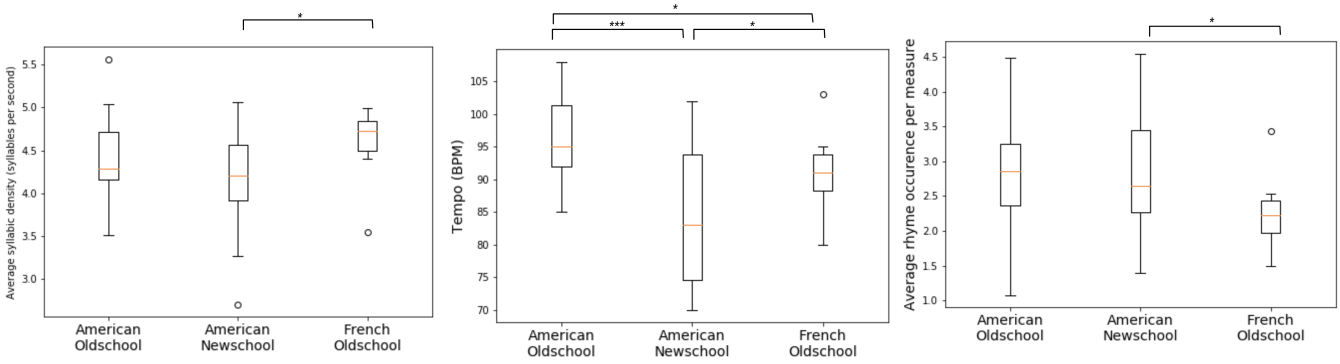
$$Pr(i|j) = \frac{\textit{total occurrence of bigram (i,j)}}{\textit{total occurrence of bigram (i,*)}}$$

where bigram *(i,*)* indicates all the bigram whose preceding syllable duration is *i*.

To make the analysis more concise, we choose 15 syllabic durations that are most frequent throughout the whole dataset. Since removing only the rare syllable duration and keeping the remaining phrase will distort the transition probability, we choose to remove the phrases that contain more than one rare duration. At the end of this step, 1.59% of American Oldschool phrases, 0.67% of American Newschool phrases, and 3.44% of French Oldschool phrases.

5. RESULTS AND DISCUSSION

The results are shown in Figure 3 - 5. Feature distributions are shown as box plots, with X indicating different groups and Y indicating the distribution of selected features. Variability of rhyme positions is shown in histogram,



(a) Distribution of syllabic density

(b) Distribution of tempo

(c) Distribution of rhyme density

Figure 3: The distribution of flow features in each group. Statistical significance is assessed using permutation tests

with X axis indicating metrical positions within 4/4 measure. One-step transition probabilities of syllable duration are shown as heat-map. For box-plot and histogram, pairwise significance based on permutation test is represented on the top of the figures.

5.1 Tempo

Figure 3a shows the distribution of average syllabic density of flows in each group. The mean syllabic density is significantly higher in French Oldschool compared to American Newschool. We observe certain tendency that the variance is smaller in French Oldschool, compared to American Oldschool and American Newschool

Figure 3b shows the distribution of tempo in each group. The mean tempo is significantly slower in American Newschool compared to American Oldschool and French Oldschool. Additionally, we conduct Levene test on the variance of tempo between Oldschool and Newschool. Between American Oldschool and American Newschool, we find that American Oldschool tempo is placed in significantly tighter range ($p = 0.016$). Between American Newschool and French Oldschool, we observe a certain tendency that the variance is smaller in French Oldschool. However in this dataset, the difference is not significant.

From these results, we can say that the strong inter-lingual difference and weak inter-genre difference in terms of syllabic tempo, On the contrary, in metrical tempo, the inter-genre difference (between American Oldschool and American Newschool) is more visible and significant.

5.2 Rhyme

Figure 3c shows the distribution of rhyme density in each group. We observe a certain tendency that the density is lower in French collection and higher in American collection. The pairwise comparison tells that the American Newschool has significantly higher mean rhyme density than American Oldschool.

Figure 4 shows the distribution of rhyme position in each group. From the shape of the distribution, we observe that the rhyme is most likely to occur during the fourth beat of measure in general. We observe that the overall entropy of rhyme position is greatest in American Oldschool, followed by American Newschool and French Oldschool. This reflects the observation that in the French Oldschool, rhyme syllables are densely positioned in the last quarter beat of the measure. Through pairwise permutation testing on entropy, we observe that the American Newschool has significantly more diverse rhyme positions than the French Oldschool.

From these results, we can say that the inter-lingual difference is more visible than the inter-genre difference, when we look into rhyme density and position diversity.

5.3 Rhythm

Figure 5 shows the transition probability of syllable duration within a flow.

In the heat map, one-step transition probabilities between syllabic duration are represented. The entry *(i, j)* indicates the probability of transition from duration i to duration j. In both axes, syllable duration is written in fractions, relative to the reference beat. The description of phrasal boundaries can be inferred from the 'start' row and the 'end' column. Since it is obvious that a 'start' phrasal boundary indicator always comes right after a 'end' phrasal boundary indicator, (end, start) transition probability is omitted from the diagram.

We observe common characteristics from all three heat maps. Firstly, (start, 1/4) entry shows that phrases are most likely to start with a note whose length is a quarter of the reference beat (a sixteenth note in 4/4 meter). Secondly, 'end' column shows that a syllable whose duration is longer than a half of the reference beat (an eighth note in 4/4 meter) is likely to become the last syllable of a phrase. Thirdly, the upper-left part of the heat map is denser than the rest of the map, provided we do not take the finishing notes into account. This means that sequences of shorter duration are more common in all groups.

At the same time, we observe differences in rhythmic transitions between different groups. First of all, the existence of diagonal lines is more prominent in the American collection than in the French collection. This expresses that in American collection, flows are more likely to consist of consecutive sequences of identical length of notes. On the other hand, from the different presence of vertical lines, we can infer different preferences for a certain dura-

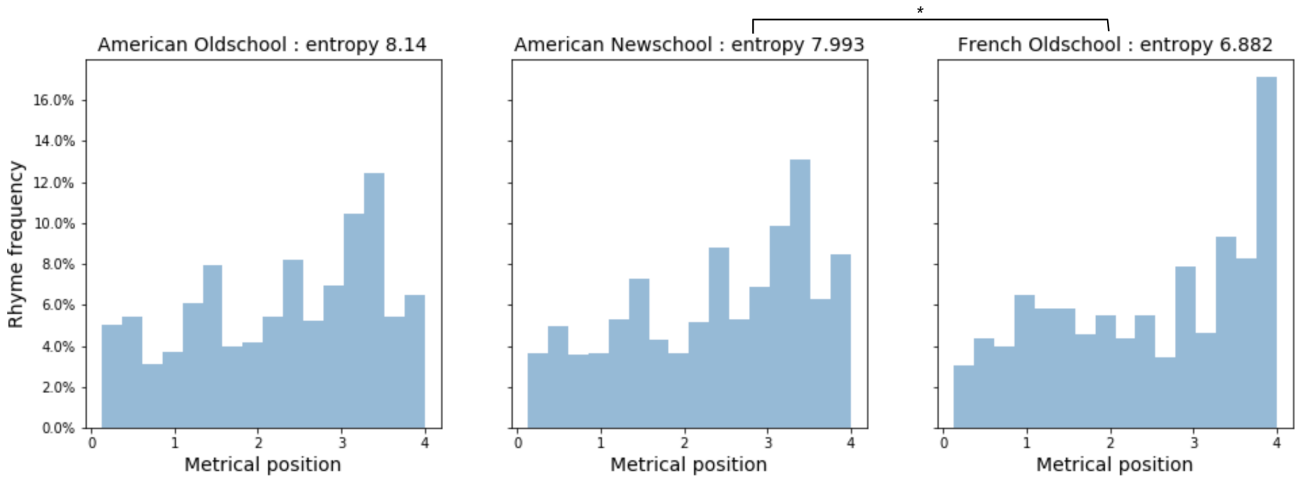


Figure 4: The distribution of rhyme positions in each group. The p-value refers to the pairwise significance of entropy. Y-axis are shared across the graphs

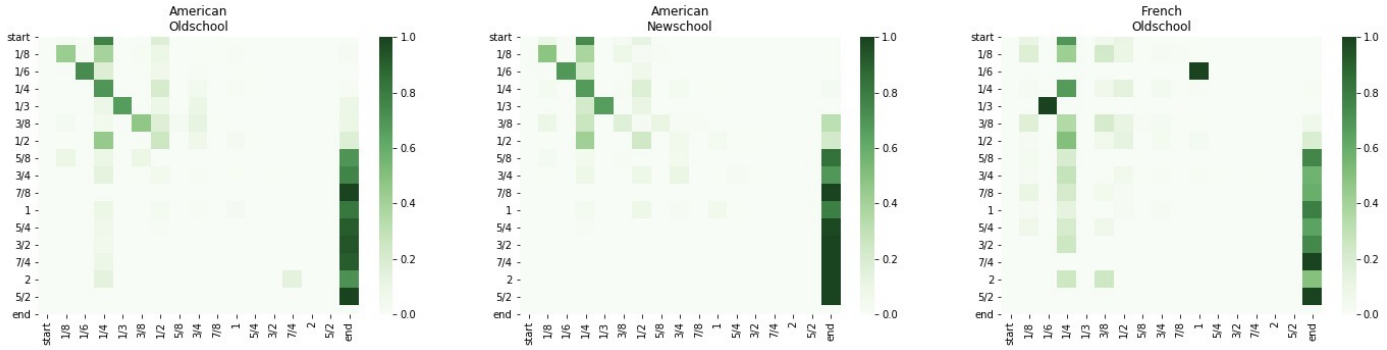


Figure 5: Representation of rhythm on one-step transition matrix. Axes represent the syllable duration written in fraction, relative to the reference beat. The transition probability of how phrase rhythm starts and ends are shown in the 'start' row and 'end' column.

tion between languages. In the American collection, vertical lines are most prominent in a quarter-beat and half-beat duration, whereas in the French collection the preference is more widely distributed, to quarter-beat, half-beat, eighth-beat and 3/8 beat. The use of triplet rhythm can be inferred from bigram sequences that contain 1/3 or 1/6 beat duration. In the American collection, sharp distribution for (1/3, 1/3) bigram and (1/6, 1/6) bigram is observed. In contrast, in the French collection, (1/3, 1/6) and (1/6, 1) bigram sequences have the sharpest distribution, which indicates that a syllable with 1/3 duration is always followed by (1/6, 1) bigram in French rap. Although the use of triplet duration is not common in all collections, we can still say that their use differs between languages.

Between American Oldschool and American Newschool, we observe that the vertical line of the quarter-beat duration is denser in the Oldschool than in the Newschool. However, the difference is less prominent than the difference between languages.

Taken together, we observe that the inter-lingual difference is greater than the inter-genre difference in rhythm analysis.

6. CONCLUSIONS AND FUTURE WORK

Taking all the results together, we can revisit the three hypotheses we made at section 2. Unlike our hypothesis that Oldschool flow is faster than Newschool flow, we observed a mixture of genre and language difference in tempo analysis. The different distribution of metrical tempo showed the strongest inter-genre difference between American Oldschool and American Newschool. In rhyme analysis, we observed that American flow uses rhyme more often and is more diverse than French flow. Finally, in rhythm analysis, we observed that the transition matrix is more similar between American Oldschool and Newschool, compared to similarity between American Oldschool and French Oldschool, which is the opposite of our third hypothesis.

In overall flow features, inter-lingual difference is more visible than inter-genre difference. Furthermore, it is worthwhile to note that the metrical tempo, the only feature that reveals strong inter-genre difference (between American Oldschool and American Newschool) is the feature that is least related to verbal aspects of rap music. Thus, we conclude that the musical characteristics of rap flows differ more greatly between languages than between genres.

Despite the findings, our study has several limitations. First of all, it lacks the corpus of French Newschool songs. Thus we couldn't determine whether the inter-lingual difference in flow is more visible than the inter-genre difference in French rap. Moreover, the accents element is ignored. Accents and stresses play an important role in any kind of musical delivery, but its importance is even greater in rap practice. Although the stress component was encoded in the original MCFlow dataset, it was not sufficient to provide concrete criteria on how to encode accents in manually encoded French dataset.

This study can be extended in several ways. Firstly, we can improve our work by having a larger corpus of French songs. It can either be achieved by adding a French Newschool corpus or having a larger French Oldschool corpus. This will increase the representative dimension of the current research and enable a better variable control. Furthermore, by applying the same methodology to corpus of other genres of rap, we can broaden our understanding on the diversity of rap flow. For example, we can apply our methodology for the Trap genre, one of the leading styles from the ten past years in Hip-Hop music, whose flows have well-defined structure and regular rhythm. We also encourage new research to cover other languages. Although we focused only on American and French rap, using our expertise in those domains, there are other languages that should be taken in account to fully understand the relationship between language and flow. Likewise, it is possible to conduct new research that discovers how different regional dialects, such as Canadian French and Moroccan French, are represented in rap flow of the same language. This will naturally lead to more detailed analysis on stresses and accents in different languages. Finally, by understanding a wider range of rap components such as lyrics in itself or the stress, future works could give a finer analysis on the rap genre in general, linking the emotions expressed in a text with the emotions associated to a certain flow.

7. REFERENCES

[1] The roots of rap classic recordings from the 1920s and 30s. "https://web.archive.org/web/20060324181308/http://www.yazoorecords.com/2018.htm". (Last accessed: 03.06.2020).

[2] Kyle Adams. On the metrical techniques of flow in rap music. 2009. https://mtosmt.org/issues/mto.09.15.5/mto.09.15.5.adams.php (Last accessed: 03.06.2020).

[3] N. Condit-Schultz. A digital corpus of rap transcriptions. *Ohio State University*, 2015.

[4] Martin Connor. The rapper's flow encyclopedia. "https://genius.com/posts/1669-The-rapper-s-flow-encyclopedia". (Last accessed: 03.06.2020).

[5] P. Edwards. *How to rap*. Chicago Review Press, Chicago, 2009.

[6] Katherine McCurdy Michael Wagnera. Poetic rhyme reflects cross-linguistic differences in information structure. *Cognition*, 117:166–175, 2010.

[7] Mitchell Ohriner. Metric ambiguity and flow in rap music: A corpus-assisted study of outkast's "mainstream" (1996). 2015. https://emusicology.org/article/view/4896/4498 (Last accessed 03.06.2020),.

[8] Catherine Tabb Powell. Rap music: An education with a beat from the street. *The Journal of Negro Education*, 60(3):245–259, 1991.

[9] SNEP. Top charts france. "http://www.chartsinfrance.net/charts/singles.php". (Last accessed: 03.06.2020).

[10] Typhlo. Les albums les plus vendus du rap français. "https://www.senscritique.com/liste/Les_albums_les_plus_vendus_du_rap_francais/1967992". (Last accessed: 03.06.2020).

[11] L. Ying. Effects of musical tempo on musicians' and non-musicians' emotional experience when listening to music. *Frontiers in Psychology*, 2009.