

# Scope of Beatrhyming: Segments or Words

Masaki Fukuda<sup>1</sup>, Kosei Kimura<sup>1</sup>, Reed Blaylock<sup>2</sup>, Seunghun J. Lee<sup>1,3</sup>

<sup>1</sup>International Christian University & <sup>2</sup>University of Southern California & <sup>3</sup>IIT Guwahati

## 1 Introduction

Beatboxing is a style of vocal musical performance in which a person produces a variety of percussive sounds, sometimes even a synthesizer or a trumpet sound, using just the human vocal tract. In this paper, the focus is on a style of beatboxing called “beatrhyming,” where a beatboxer sings a song with lyrics while producing beatboxing sounds at the same time. Beatrhyming is unique from both beatboxing and linguistic perspectives since it is a performance where sounds of speech and of beatboxing co-exist. Beatrhyming therefore provides new opportunities to explore the flexibility of speech. For example, Blaylock & Phoolsombat (2019) found that in beatrhyming, beatboxing sounds replace speech sounds in the lyrics of a song by matching place of articulation, such as a labial Kick Drum (IPA: [pʼ]) replacing a labial speech plosive in the word “baby” /bebi/ → [pʼebi]. However, their case study was limited to only one song and did not address whether whole words can be replaced in beatboxing. This paper explores the way three different beatboxers combine beatboxing and speech in beatrhyming.

## 2 Method

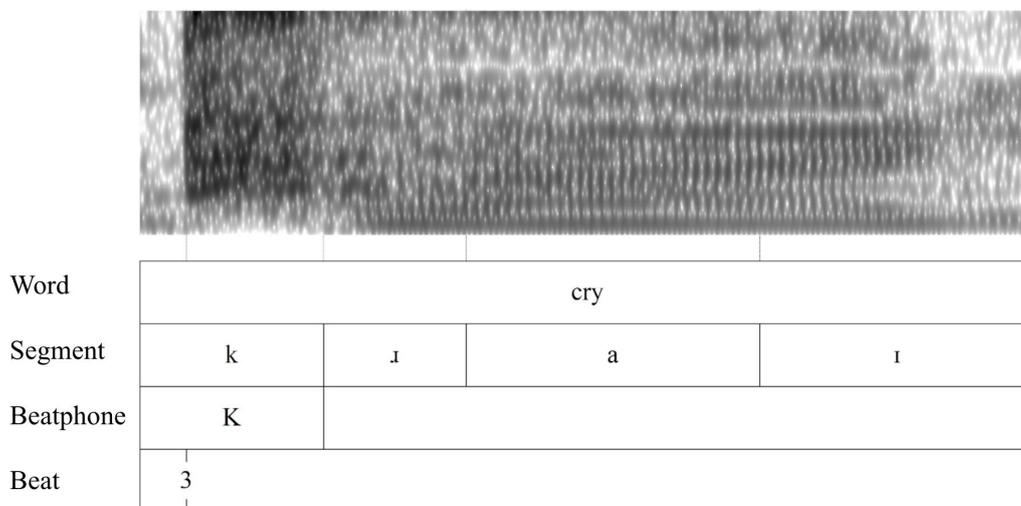
The data come from three different beatrhyming performances found on YouTube: a cover of Childish Gambino’s “Redbone” performed by Gene (The Beatbox House, 2017, <https://youtu.be/PItO1YQ89Hg>), D-Low’s performance in the “Grand Beatbox SHOWCASE Battle 2019” (Swissbeatbox, 2019, <https://youtu.be/DFk6aVSWo7s>), and “Grand Beatbox Battle 2021 World League Solo Wildcard” performed by Heartzel (Heartzel, 2021, <https://youtu.be/iKXVQI9E9DQ>). These three performances were selected because they featured multiple relatively long speech phrases and sentences (as opposed to some beatboxing performances which include words but only sporadically).

Each piece was manually labeled along four tiers as depicted in Figure 1: Word, Segment, Beatphone, and Beat. Labels on the Word tier orthographically denoted individual words produced by the beatboxer. Annotations along the Segment tier marked the acoustic intervals for the segments of each word on the Word tier, and annotations along the Beatphone tier marked the acoustic intervals for each beatboxing sound. A defining characteristic of beatrhyming is the overlap or interleaving of beatboxing and speech sounds, often resulting in one or more speech sounds appearing to be replaced by a beatboxing sound. In cases of apparent replacement, labels were made for both the speech segment(s) typically associated with the word being produced and the beatboxing sound(s) that replaced them. Figure 1 shows an example of this with /k/ of “cry” being replaced by {K}. Finally, the Beat tier was labeled with points that marked the musical beat associated with each beatboxing sound, aligned to the onset of the acoustic release bursts of stops and affricates or to the onset of regular vibration for trills. Labeling was conducted in two phases: in the first phase, the first and second authors (both expert beatboxers) made annotations in consultation with each other and revised those annotations based on discussions involving all four authors; in the second phase, the third author (a phonetician but novice beatboxer) made minor revisions.

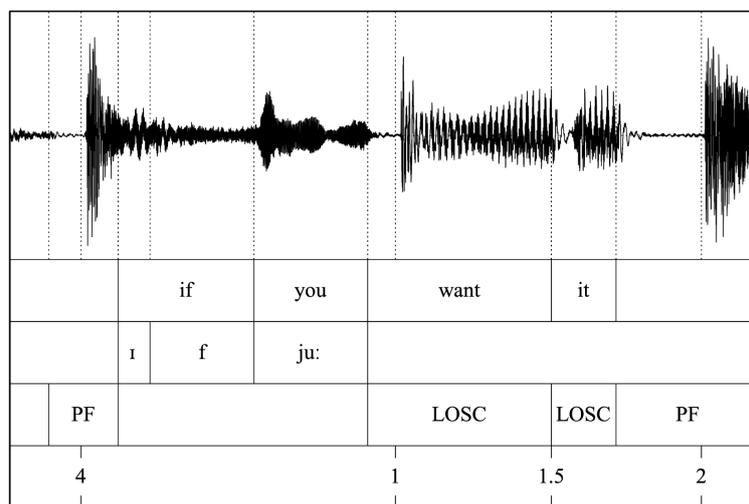
It can sometimes be difficult to determine whether a given sound produced in a beatrhyming performance is a speech sound or a beatboxing sound. In particular, Kick Drums [pʼ] and Closed Hi-Hats [tsʼ] sometimes resemble spoken English labial and alveolar stops, respectively. This is advantageous for the beatrhymers replacing speech sounds with beatboxing sounds to simultaneously perform an aesthetically pleasing beatboxing pattern and convey a linguistic message, but it can be challenging for the analyst to decide whether a sound is a speech sound or a beatboxing sound replacing a speech sound. In cases where a word being produced was known and the place of articulation of a sound in the word did not match its expected place of articulation, then that sound was labeled as being replaced by a beatboxing sound (ex. the /t/ in “tell” replaced with a K Snare {^K}). When there were no place of articulation discrepancies, identification of a sound as a speech sound or a beatboxing sound was done by the beatboxing annotators’ intuitions, reference to the Beat tier (as some beats are commonly associated with the production of a beatboxing sound), visual assessment of burst release duration (i.e., Kick Drums and Closed Hi-Hats tend to have longer release bursts than voiced labial and alveolar

stops), and impressionistic comparisons between the sound in question and other sounds in the same beat pattern that had previously been definitively classified as a speech sound or a beatboxing sound.

Two types of replacements were counted for this study: segment replacements and word replacements. A segment replacement was counted if a beatboxing sound had complete temporal overlap with a single intended speech segment, as in Figure 1 with {K} replacing [k] in “cry.” A word replacement was counted if all the intended segments of a word were produced as a beatboxing sound, as in Figure 2 where the words “want” and “it” are entirely produced as Lip Oscillations.



**Figure 1:** Praat annotation of the word “cry” beathrmed by D-Low. The {K} on the Beatphone tier temporally overlaps with the [k] on the Segment tier, denoting that the [k] in [kɪaɪ] has been completely replaced by a K Snare {K}, resulting in the pronunciation of “cry” as [k<sup>^</sup>ɪ↓aɪ].



**Figure 2:** Praat annotation of the phrase “if you want it” beathrmed by Gene. The words “want” and “it” are fully replaced by Lip Oscillations (LOSC), a type of labial trill. The brief period of silence and subsequent burst at the beginning of the Lip Oscillation for “want” indicate that the trill is initiated with a Kick Drum [p’], which is very common for labial trills in beatboxing.

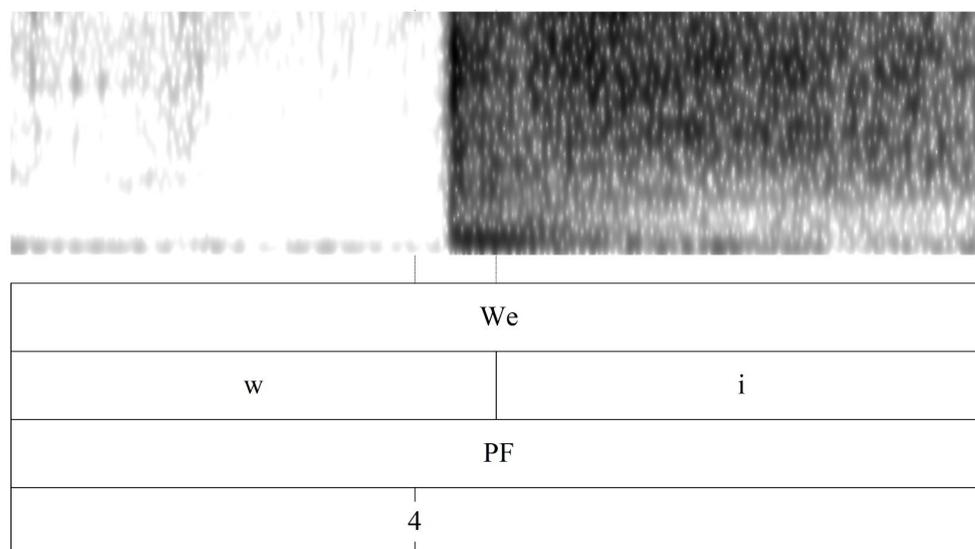
### 3 Results

**3.1 Beatboxing sounds** The three beatboxers each used a subset of the sounds shown in Table 1. The Standard Beatbox Notation for each sound is provided when known (Splinter & Tyte, 2002), as is a phonetic description of the sound and IPA notation based on the phonetic description. Note that the ingressive symbol [↓] and ventricular phonation symbol [!] are drawn from the Extensions to the IPA for Disordered Speech and the Voice Quality Symbols of Ball, Esling & Dickson (1995). The ingressive symbol is typically meant to indicate pulmonic ingressive airflow as used by the Inward K Snare {<sup>^</sup>K}, but we also use it here to indicate the lingual ingressive airflow of the Liproll for which no IPA symbol was immediately apparent. There was some disagreement about whether some replacements were performed with Throat Bass or Lip Bass (a type of bilabial



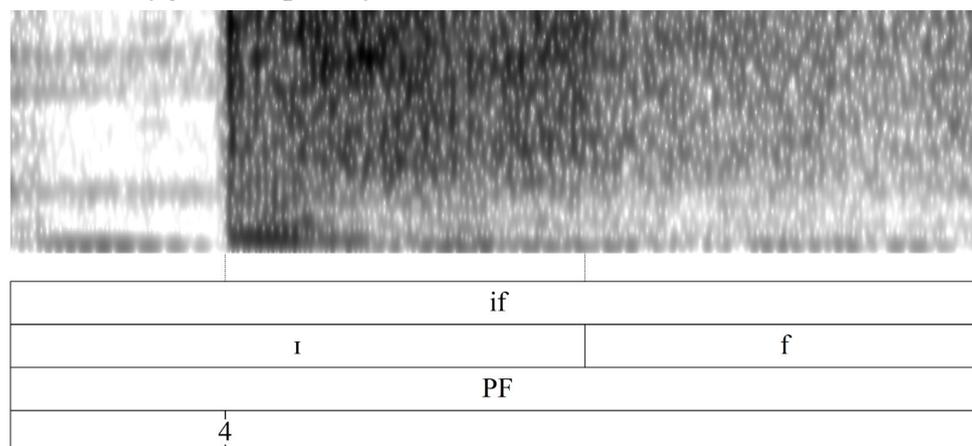
## 4 Discussion

During data analysis, it was sometimes controversial whether PF Snares were performing segment-level replacements or word-level replacements. However, the PF Snare sometimes clearly replaces a whole word, as shown in the Figure 3 with the replacement of “we” [wi] by a PF Snare. This appears to be done by extending the duration of the frication of the PF Snare.



**Figure 3:** A PF Snare performing word-level replacement, replacing both segments of “we” [wi].

Not all cases are so clear. Figure 5 shows another example in which a PF Snare was annotated as replacing the word “if” [ɪf]. However, it was difficult to determine whether the PF snare replaces the whole word “if” or it only replaces the segment /ɪ/. Judging from the spectrogram, there is no vowel observed, therefore it is clear that the PF snare at least replaces the segment /ɪ/. However, the overall spectral intensity dissipates about half-way through the fricative, and it is unclear whether the intensity shift marks a change from the PF Snare to the [f] of “if” or the PF Snare simply becoming a bit quieter.



**Figure 4:** A PF Snare performing either word-level or segment-level replacement. The TextGrid in this figure indicates word-level replacement, with the PF Snare replacing both segments of “if” [ɪf].

Phonetically, affricates are complex segments that can be decomposed into a plosive and a fricative. The plosive part is short with a transient noise (the release burst), while the fricative portion has a sustained frication noise. Even so, affricates phonologically form a single segment. We suggest that this phonetic, but not phonological, nature of affricates is reflected in the sound replacement patterns by beatboxing sounds.

Segment replacements (Closed Hi-Hat, (Inward) K Snare) are possible because affricates are produced with transient noise, and word replacements (PF Snare) occur as affricates have sustained frication noise. This dichotomy predicts that sounds with transient noise are likely to participate in segment replacement, which is the case as Kick Drum replaces stop segments. Reversely, sustained sounds are likely to show word replacements;

trills and phonation create sustained sounds. These sounds with sustained duration exhibit word replacements. Since affricates have both properties, they may be used more flexibly in beathyming for either segment-level or word-level replacements. Clearly further studies are needed but our data suggests that phonetic and phonological knowledge of a speech segment is actively used by beatboxers when they produce beathyming patterns.

## 5 Conclusion

This result helps us understand that continuous sounds such as trills and phonations tend to replace whole words, and short sounds such as stops or affricates usually replace segments. The existence of word-level replacements in beathyming is a new finding for beathyming research, and the pattern of affricates replacements either words or segments hints at a connection or parallel between a beathymer's beatboxing and their phonological knowledge. Beathyming provides exciting new ways to explore the relationship between a speaker's phonological knowledge and non-linguistic behaviors like beatboxing.

## References

- Blaylock, R., & Phoosombat, R. 2019. Beatrhyming probes the nature of the interface between phonology and beatboxing. *The Journal of the Acoustical Society of America*, 146(4), 3081 – 3081. <https://doi.org/10.1121/1.5137696>
- The Beatbox House. 2017. Gene | Redbone (Beatbox Cover) [Video]. YouTube. <https://youtu.be/PltO1YQ89Hg>. Retrieved August 4th, 2021.
- Heartzel. 2021. HEARTZEL | GBB 2021 World League Solo Wildcard | 11th place! [Video]. YouTube. <https://youtu.be/iKXVQ19E9DQ>. Retrieved August 4th, 2021.
- Swissbeatbox. 2019. D-LOW | Grand Beatbox Battle Champion 2019 Compilation [Video]. YouTube. <https://youtu.be/DFk6aVSWo7s>. Retrieved August 4th, 2021.
- Splinter, M. and TyTe, G. 2002. Standard Beatbox Notation (SBN). <https://humanbeatbox.com/articles/standard-beatbox-notation-sbn/>. Retrieved December 8, 2019.
- Ball, Esling & Dickson. 1995. "The VoQS System for the Transcription of Voice Quality", *Journal of the International Phonetic Association*, 25.02, p. 71 – 80.

## Acknowledgments

This research is funded by Kakenhi (C) 20K00578, and The ILCAA joint research project "Phonetic typology from cross-linguistic perspectives (PhonTyp)" (2021-2023). We thank the audience members of AJL6 for their comments and suggestions.