Evidence for phonological features and phonotactics in beatboxing vocal percussion

Devon Guinn (Harvard University) and Aleksei Nazarov (University of Huddersfield)

Beatboxing (Stowell and Plumbley 2008), a vocal technique originating in the imitation of percussion in early hip-hop, uses the same articulatory apparatus as spoken language and shows major overlap with linguistic sounds. However, it is not *a priori* clear whether consonantal articulations in beatboxing have the same phonological representations as natural language. For instance, it could be that beatboxing only uses raw phonetic specifications presupposing no phonological labeling (unlike many exemplar models for linguistic phonology – e.g., Pierrehumbert 2001), e.g., specific articulatory instructions (see, for instance, Hoole & Pouplier 2015) with a specified tolerance of error. We show that beatboxing consonants exhibit natural class behavior, suggesting that these sounds are labeled with something similar to linguistic phonological features. In addition, we show that these natural classes participate in phonotactic constraints that operate similarly to natural language phonotactic constraints. These findings suggest that the system of beatboxing uses phonological-like representations, and thus may be able to inform phonological theory much in the same way as data from traditional poetry (e.g., Ryan 2011) and hip-hop (Katz 2015) have informed it before.

Arguments for natural classes from morphophonology (e.g., Chomsky & Halle 1968) are impossible in beatboxing because it has no semantics or (linguistic) morphology. Instead, our argument comes from variable realization of repeated rhythmic templates, as illustrated in (1). Beatboxers have intuitions about the (im)possibility of certain templates. Consulting this "beatboxer's intuition", we found that templates like (1a) are possible, where the 1st and 3rd sound, both labial plosives, may be produced egressively (standardly as an ejective, [pʼ]) or ingressively (standardly as an implosive, [pʼ↓]). The equally possible template in (1c) exhibits wider variation: the 1st and 3rd sound may be any voiceless labial sound that is not a fricative, nasal, or approximant: a bilabial stop or trill with any airstream mechanism and direction.

Crucially, it is impossible for a sound in a template to vary between [p$^{(ʼ)}$ ~ p$^{(ʼ)}$↓ ~ ʙ̥$^{(ʔ)}$], excluding ingressive trills, as in (1b). This is surprising: these 3 sounds form a more constrained articulatory space than those in the attested (1c): [p$^{(ʼ)}$↓] and [ʙ̥$^{(ʔ)}$] are one major articulatory change away from [p$^{(ʼ)}$] (in airstream direction and constriction degree, respectively), while ingressive [ʙ̥$^{(ʔ)}$↓] has two major changes w.r.t. [p$^{(ʼ)}$] (airstream direction + constriction degree).

C 𝄆 ♩ ♩ 𝅘𝅥𝅯 ♩ 𝄇

(1) a. /P pfʼ P pfʼ/ → [p$^{(ʼ)}$ pfʼ p$^{(ʼ)}$ pfʼ ~ p$^{(ʼ)}$↓ pfʼ p$^{(ʼ)}$↓ pfʼ]

/P/ = /+labial, –delayed release, –nas, –vce, (–approximant,) –sonorant/

b. (impossible) → *{ [p$^{(ʼ)}$ pfʼ p$^{(ʼ)}$ pfʼ ~ p$^{(ʼ)}$↓ pfʼ p$^{(ʼ)}$↓ pfʼ ~ ʙ̥$^{(ʔ)}$ pfʼ ʙ̥$^{(ʔ)}$ pfʼ] }

(no feature description for [p$^{(ʼ)}$ ~ p$^{(ʼ)}$↓ ~ ʙ̥$^{(ʔ)}$])

c. /β pfʼ β pfʼ/ → [p$^{(ʼ)}$ pfʼ p$^{(ʼ)}$ pfʼ ~ p$^{(ʼ)}$↓ pfʼ p$^{(ʼ)}$↓ pfʼ ~ ʙ̥$^{(ʔ)}$ pfʼ ʙ̥$^{(ʔ)}$ pfʼ ~ ʙ̥$^{(ʔ)}$↓ pfʼ ʙ̥$^{(ʔ)}$↓ pfʼ]

/β/ = /+labial, –delayed release, –nas, –vce, –approximant, 0 sonorant/

If these templates had no categorical representations – e.g., if the first beat contained an instruction for [p$^{(ʼ)}$] with a set tolerance of error – we would expect that the more articulatorily constrained variation in (1b) were at least as acceptable as the less constrained variation in (1c). Instead, the fact that only (1a) and (1c) are possible suggests that beatboxing mental representations contain a categorical feature like [±sonorant]: as shown in (1a), specifying [–sonorant] means no trills are included, while, as shown in (1c), underspecifying [sonorant] means all voiceless labial trills compatible are included, both ingressive and egressive.

Crucially, there is no feature combination that would specify the impossible variation in (1b). Similar substitution data motivate place and laryngeal features (e.g., [±labial], [±cons. glottis]).

In addition to this, we found that natural classes with respect to place participate in categorical phonotactic constraints. Specifically, accented positions cannot be filled by coronal consonants, as illustrated in example (2). In this example, different accentuation of the same string of consonants yields opposite acceptability; accented consonants are shown with <x> on a rudimentary metrical grid above each sequence. In (2ab), the first shaded consonant is a coronal [ts'], followed by non-coronal [p']. When the accent is on the second sound and [ts'] is unaccented, as in (2a), the sequence is well-formed, but when the accent is on the first sound, [ts'], as in (2b), the sequence is unacceptable. In (2cd), the shaded [ts'] and [p'] swap places. Now, the same accentuation patterns yield opposite acceptability: the sequence in (2c), which has an accent on the second shaded sound, [ts'], is unacceptable, but when the same [ts'] is unaccented, as in (2d), the sequence is well-formed.

(2)

C ‖: ♪ ♪ ♪ ♪ ♪ ♪ ♪ ♪ :‖

a. x x x x
[p' ts' pf' ts' ts' p' pf' ts']

b. x x x x
*[p' ts' pf' ts' ts' p' pf' ts']

c. x x x x
*[p' ts' pf' ts' p' ts' pf' ts']

d. x x x x
[p' ts' pf' ts' p' ts' pf' ts']

The judgments in (2) hold when any coronal consonant used in beatboxing (namely, [s, t, ts, tʃ] in their ingressive and egressive, glottalic and velaric forms) is used instead of [ts']. At the same time, most non-coronal sounds are indeed allowed in accented positions, except for the fricative [f]. This provides evidence for a phonotactic constraint against coronal consonants in accented positions, along the lines of that in (3). A similar constraint may also be posited for fricatives (coronal and non-coronal): they, too, are banned in accented position.

(3) Phonological constraint against accented coronals

* x
[+cor]

These initial results indicate that the mental representation of beatboxing contains categorical features and phonotactic constraints like linguistic phonological grammars do. This suggests that patterns unique to beatboxing might bear on what is (dis)allowed by phonological theory, making this musical genre a potential new source of evidence for phonologists.

References Chomsky, N. & M. Halle. 1968. *The Sound Pattern of English*. Harper & Row. • Hoole, P. & M. Pouplier. 2015. Interarticulatory coordination - speech sounds. In M. Redford (ed.) *The Handbook of Speech Production*. Wiley. 133-57. • Katz, Jonah. 2015. Hip-hop rhymes reiterate phonological typology. *Lingua* 160:54-73. • Pierrehumbert, J. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (eds.) *Frequency effects and the emergence of lexical structure*. John Benjamins. 137-57. • Ryan, Kevin. 2011. *Gradient Weight in Phonology*. PhD dissertation, UCLA. • Stowell, D. & M.D. Plumbley. 2008. *Characteristics of the beatboxing vocal style.* Tech. Rep. C4DM-TR-08-01, Dept. of Electronic Engineering, Queen Mary University of London.