

# Characteristics of the beatboxing vocal style

Technical report C4DM-TR-08-01

Dan Stowell and Mark D. Plumbley

Centre for Digital Music

Department of Electronic Engineering

Queen Mary, University of London

`dan.stowell@elec.qmul.ac.uk`

19th February 2008

## 1 Introduction

Beatboxing is a tradition of vocal percussion which originates in 1980s hip-hop, and is closely connected with hip-hop culture. It involves the vocal imitation of drum machines as well as drums and other percussion, and typically also the simultaneous imitation of basslines, melodies, and vocals, to create an illusion of polyphonic music. It may be performed *a capella* or with amplification. In this report we describe some characteristics of the beatboxing vocal performance style, as relevant for music signal processing and related fields. In particular we focus on aspects of beatboxing which are different from other vocal styles or from spoken language.

Beatboxing developed well outside academia, and separate from the vocal styles commonly studied by universities and conservatories, and so there is (to our knowledge) very little scholarly work on the topic, either its history or its current practice. Beatboxing is mentioned in popular histories of the hip-hop movement, although rarely in detail. An undergraduate thesis looks at phonetic aspects of some beatboxing sounds [Lederer, 2005]. Some technical work is inspired by beatboxing to create (e.g.) a voice-controlled drum-machine [Hazan, 2005a,b, Kapur et al., 2004, Sinyor et al., 2005], although these authors don't make explicit whether their work has been developed in contact with practising beatboxers.

In the following we describe characteristics of beatboxing as contrasted against better-documented traditions such as popular singing [Soto-Morettini, 2006] or classical singing [Mabry, 2002]. Because of the relative scarcity of literature, many of the observations come from the first author's experiences and observations: both as a participant in beatboxing communities in the UK and online, and during user studies involving beatboxers as part of the first author's PhD study.

We describe certain sounds narratively as well as in International Phonetic Alphabet (IPA) notation [International Phonetic Association, 1999] (see also [Fukui, 2003]), which will be demarcated by slashes // . The IPA representation may be approximate, since the notation is not designed to accommodate easily the non-linguistic and "extended technique" sounds we discuss.

## 2 Extended vocal technique

Perhaps the most fundamental distinction between the sounds produced while beatboxing and those produced during most other vocal traditions arises from beatboxing's primary aim to create *convincing impersonations* of drum tracks. (Contrast this against vocal percussion traditions such as jazz *scat* singing or indian *bol*, in which percussive rhythms are imitated, but there is no aim to disguise the vocal origin of the sounds.) This aim leads beatboxers to do two

things: (1) employ a wide palette of vocal techniques to produce the desired timbres; and (2) suppress some of the linguistic cues that would make clear to an audience that the source is a single human voice.

The extended vocal techniques used are many and varied, and vary according to the performer. Many techniques are refinements of standard linguistic vowel and consonant sounds, while some involve sounds that are rarely if at all employed in natural languages. We do not aim to describe all common techniques here, but we will discuss some relatively general aspects of vocal technique which have a noticeable effect on the sound produced.

## 2.1 Non-syllabic patterns

The musical sounds which beatboxers imitate may not sound much like conventional vocal utterances. Therefore the vowel-consonant alternation which is typical of most use of voice may not be entirely suitable for producing a close auditory match. Instead, beatboxers learn to produce sounds to match the sound patterns they aim to replicate, attempting to overcome linguistic patternings. Since human listeners are known to use linguistic sound patterns as one cue to understanding a spoken voice [Shannon et al., 1995], it seems likely that avoiding such patterns may help maintain the illusion of non-voice sound.

As mentioned above, vocal traditions such as *scat* or *bol* do not aim to disguise the vocal origin of the sounds. Hence in those traditions, patterns are often built up using syllable sounds which do not stray far from the performers' languages.

## 2.2 Use of inhaled sounds

In most singing and spoken language, the vast majority of sounds are produced during exhalation. (Many languages do allow a minor linguistic role for inhaled phonation [Ladefoged and Maddieson, 1997]. Some vocal performance traditions feature inhaled sounds, e.g. Inuit throat-singing games [Nattiez, 2008].)

A notable characteristic of beatboxing is the widespread use of inhaled sounds. We propose that this has two main motivations. Firstly it enables a continuous flow of sounds, which both allows for continuous drum patterns and also helps maintain the auditory illusion of the sounds being imitated (since

the sound and the pause associated with an ordinary intake of breath are avoided). Secondly it allows for the production of certain sounds which cannot be produced equally well during exhaling. A commonly-used example is the “inward clap snare” /k<sup>l</sup>/ <sup>1</sup>.

Inhaled sounds are most commonly percussive. Although it is possible to phonate while breathing in, the production of pitched notes while inhaling does not seem to be used much at all by beatboxers.

Although some sounds may be specifically produced using inward breath, there are many sounds which beatboxers seem often to be able to produce in either direction, such as the “closed hi-hat” sound /t<sup>ɾ</sup>/ (outward) or /t<sup>ɿ</sup>/ (inward). This allows some degree of independence between the breathing patterns and the rhythm patterns.

## 2.3 Vocal modes/qualities

Laver [1980] provides the classic phonetician's description of the different voice qualities or “phonatory settings” that an individual can produce, including falsetto, creaky voice, harsh voice, breathy voice, and ventricular voice. (The term “modal voice” is also employed, to refer to the most common vocal quality against which these others are to be distinguished.) These qualities may be consciously manipulated by a speaker, may be part of linguistic distinctions between vowels, or may be indicative of vocal pathology. In study of the singing voice, too, different vocal modes are distinguished [Soto-Moretini, 2006], including head voice, chest voice, belt, twangy voice, growl, breathy voice and creaky voice. Note the (incomplete) overlap between the categories used by the two communities.

Beatboxers make use of different vocal qualities to produce specific sounds. For example, growl/ventricular voice may be used to produce a bass tone, and falsetto is used as a component of some sounds, e.g. vocal scratch, “synth kick”. In these cases the vocal qualities are employed for their timbral effects, not (as may occur in language) to convey meaning or emotional state.

Some beatboxing techniques involve the alternation between voice qualities. If multiple streams

---

<sup>1</sup>[http://www.humanbeatbox.com/inward\\_snares](http://www.humanbeatbox.com/inward_snares)

are being woven into a single beat pattern, this can involve rapid alternation between (e.g.) beats performed using modal voice, “vocals” or sound effects performed in falsetto, and basslines performed in growl/ventricular voice. The alternation between voice qualities can emphasise the separation of these streams and perhaps contribute to the illusion of polyphony.

## 2.4 Trills / rolls / buzzes

Beatboxers tend to use a variety of trills to produce oscillatory sounds. (Here we use the term “trill” in its phonetic sense, as an oscillation produced by a repeated blocking and unblocking of the airstream; *not* in the musical sense of a rapid alternation between pitches.) The IPA explicitly recognises three trill types:

- /r/ (alveolar trill or “rolled R”)
- /β/ (voiced bilabial trill)
- /R/ (uvular trill)

These have a role in beatboxing, as do others: trills involving the palate, inward-breathed trills and click-trills.

The frequency of vocal trills can vary from subsonic rates (e.g. 20–30 Hz) to low but audible pitches (e.g. 100 Hz) [Ladefoged and Maddieson, 1997, chapter 7]. This leads to trills being employed in two distinct ways: (1) for rapidly-repeated sounds such as drum-rolls or “dalek” sound (the gargling effect of uvular trill); and (2) for pitched sounds, particularly bass sounds. In the latter category, bilabial trill (“lip buzz”) is most commonly used, but palatal trills and inward uvular trills (“snore bass”) are also used.

Notably, beatboxers improve the resonant tone of pitched trills (particularly /β/) by matching the trill frequency with the frequency of voicing. This requires practice (to be able to modify lip tension suitably), but the matched resonance can produce a very strong bass tone, qualitatively different from an ordinary voiced bilabial trill.

A relatively common technique is the “click roll”, which produces the sound of a few lateral clicks in quick succession: /l̥l̥l̥l̥/. This is produced by the

tongue and palate and does not require the intake or exhaling of air, meaning (as with other click-type sounds) that beatboxers can produce the sound simultaneously with breathing in or with humming. (There exist click-roll variants produced using inhaled or exhaled breath.)

Although trilling is one way to produce drum-roll sounds, beatboxers do also use fast alternation of sounds as an alternative strategy to produce rapidly-repeated sounds, e.g. /b̥d̥b̥d̥b̥d̥/ for kicks or /t̥f̥t̥f̥t̥f̥/ for hi-hats <sup>2</sup>.

## 3 Close-mic technique

Beatboxing may be performed *a capella* or with a microphone and amplification. In the latter case, many beatboxers adopt a “close-mic” technique: while standard dynamic microphones are designed to be used at a distance of around 15–20 centimetres from the mouth for a “natural” sound quality [Shure Inc., 2006], beatboxers typically use a standard dynamic vocal mic but positioned around one or two centimetres from the mouth<sup>3</sup>. This is to exploit the response characteristics of the microphone at close range, typically creating a bassier sound [Shure Inc., 2006]. The performer may also cup the microphone with one or both hands to modulate the acoustic response.

For some sound qualities or effects the microphone may be positioned against the throat or the nose. Against the throat, a muffled “low-pass filter” effect can be produced.

Close-mic techniques alter the role of the microphone, from being a “transparent” tool for capturing sound to being a part of the “instrument”. There is an analogy between the development of these techniques, and the developments following the invention of the electric guitar, when overdrive and distortion sounds (produced by nonlinearities in guitar amplifiers) came to be interpreted, not as deviations from high fidelity, but as specific sound effects.

<sup>2</sup><http://www.humanbeatbox.com/rolls>

<sup>3</sup>[http://www.humanbeatbox.com/techniques/p2\\_articleid/128](http://www.humanbeatbox.com/techniques/p2_articleid/128)

## 4 Conclusion

Beatboxing is a relatively recently-developed performance style involving some distinct performance techniques which affect the nature of the audio stream, compared against the audio produced in most other vocal performance styles. The use of non-syllabic patterns and the role of inhaled sounds typically leads to an audio stream in which language-like patterns are suppressed, which we argue may facilitate the illusion of a non-vocal sound source(s). These and other extended vocal techniques are employed to provide a diverse sound palette. Close-mic techniques are used explicitly to modify the characteristics of the sound.

In this report we have documented aspects of these performance techniques, and hope to have provided details to illuminate how the performance style may affect the nature of the recorded sound, as contrasted against other vocal musical performance styles.

## References

- R. Fukui. *TIPA Manual*, 2003. URL <http://www.ctan.org/tex-archive/fonts/tipa/tipaman.pdf>.
- A. Hazan. Towards automatic transcription of expressive oral percussive performances. In *Proceedings of the 10th international conference on Intelligent User Interfaces (IUI '05)*, pages 296–298, New York, NY, USA, 2005a. ACM Press. ISBN 1-58113-894-6. doi: 10.1145/1040830.1040904.
- A. Hazan. Billaboop: Real-time voice-driven drum generator. In *Proceedings of the 118th Audio Engineering Society Convention (AES 118)*, May 2005b.
- International Phonetic Association. *Handbook of the International Phonetic Association: a guide to the use of the international phonetic alphabet*. Cambridge University Press, 1999.
- A. Kapur, M. Benning, and G. Tzanetakis. Query-by-beat-boxing: Music retrieval for the DJ. In *Proceedings of the Fifth International Conference on Music Information Retrieval*, 2004.
- P. Ladefoged and I. Maddieson. *The Sounds of the World's Languages*. Blackwell, 1997.
- J. Laver. *The Phonetic Description of Voice Quality*. Cambridge Studies in Linguistics. Cambridge University Press, 1980. URL [http://www.ling.mq.edu.au/ling/units/sph302/papers/laver\\_1980\\_phonation.pdf](http://www.ling.mq.edu.au/ling/units/sph302/papers/laver_1980_phonation.pdf).
- K. Lederer. The phonetics of beatboxing (undergraduate degree dissertation), 2005. URL <http://www.humanbeatbox.com/phonetics>.
- S. Mabry. *Exploring Twentieth-Century Vocal Music: A Practical Guide to Innovations in Performance and Repertoire*. Oxford University Press, USA, 2002.
- J.-J. Nattiez. *The Canadian Encyclopedia*, chapter Inuit vocal games. Historica Foundation, 2008. URL <http://www.thecanadianencyclopedia.com/index.cfm?PgNm=TCE&Params=U1ARTU0001711>.
- R. V. Shannon, F.-G. Zeng, V. Kamath, J. Wygon-ski, and M. Ekelid. Speech recognition with primarily temporal cues. *Science*, 270(5234):303–304, Oct 1995. doi: 10.1126/science.270.5234.303. URL <http://www.sciencemag.org/cgi/content/abstract/270/5234/303>.
- Shure Inc. *Shure SM58 user guide*, 2006. URL [http://www.shure.com/stellent/groups/public/@gms\\_gmi\\_web\\_ug/documents/web\\_resource/us\\_pro\\_sm58\\_ug.pdf](http://www.shure.com/stellent/groups/public/@gms_gmi_web_ug/documents/web_resource/us_pro_sm58_ug.pdf).
- E. Sinyor, C. McKay, R. Fiebrink, D. McEnnis, and I. Fujinaga. Beatbox classification using ACE. In *Proceedings of the International Conference on Music Information Retrieval*, 2005.
- D. Soto-Morettini. *Popular Singing: A Practical Guide To: Pop, Jazz, Blues, Rock, Country and Gospel*. A & C Black, 2006. ISBN 978-0713672664.