



**HAL**  
open science

## Humming beatboxing: the vocal orchestra within

Annalisa Paroni, H el ene Loevenbruck, Pierre Baraduc, Christophe Savariaux,  
Pascale Calabrese, Nathalie Henrich Bernardoni

► **To cite this version:**

Annalisa Paroni, H el ene Loevenbruck, Pierre Baraduc, Christophe Savariaux, Pascale Calabrese, et al.. Humming beatboxing: the vocal orchestra within. MAVEBA 2021 - 12th International Workshop Models and Analysis of Vocal Emissions for Biomedical Applications, Universita Degli Studi Firenze, Dec 2021, Florence, Italy. hal-03510719

**HAL Id: hal-03510719**

**<https://hal.science/hal-03510719v1>**

Submitted on 4 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

# HUMMING BEATBOXING: THE VOCAL ORCHESTRA WITHIN

Annalisa Paroni<sup>1</sup>, H el ene L oevenbruck<sup>2</sup>, Pierre Baraduc<sup>1</sup>, Christophe Savariaux<sup>1</sup>,  
Pascale Calabr ese<sup>3</sup>, and Nathalie Henrich Bernardoni<sup>1</sup>

<sup>1</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, F-38000 Grenoble

<sup>2</sup>Univ. Grenoble Alpes, Univ. Savoie Mont-Blanc, CNRS, LPNC, F-38000 Grenoble

<sup>3</sup>CNRS, Grenoble INP, TIMC-IMAG, F-38000 Grenoble

annalisa.paroni@gipsa-lab.fr, Helene.Loevenbruck@univ-grenoble-alpes.fr,  
pierre.baraduc@gipsa-lab.fr, christophe.savariaux@gipsa-lab.fr,  
Pascale.Calabrese@univ-grenoble-alpes.fr, nathalie.henrich@gipsa-lab.fr

**Abstract:** Humming beatboxing is a technique used by beatboxers to give the impression of producing multiple sounds synchronously. How this is achieved and what the differences are with regular beatboxing remains mostly unexplored from a scientific standpoint. Four beatboxers were recorded. Electromagnetic articulography was combined with acoustic, electroglottographic, and breathing measurements. The articulatory and breathing behaviors of three boxemes (kick, hi-hat, rimshot) were compared between a regular and a humming realization. When produced as regular beatboxing, the trajectories of the tongue were consistent with a glottalic initiation mechanism and breathing behavior was related to the acoustic outcome. In contrast, for humming sounds the breathing behavior was dissociated from articulation and acoustic outcome, suggesting that the initiation mechanism took place within the oral cavity and the more posterior portion of the vocal tract did not participate in the production of those sounds. Articulatory trajectories were consistent with a closure held at the posterior region of the oral cavity. This suggests that in the humming technique the vocal tract is divided into two sections: the oral cavity functions on its own to produce the rhythmic line, while the melodic line is produced in the laryngeal or pharyngeal spaces and propagates through the nasal cavities. **Keywords:** Human beatboxing, humming beatboxing, humming

## I. INTRODUCTION

Human beatboxing (HBB) is an emerging and rapidly evolving vocal art that relies on the human vocal instrument to produce all kinds of sounds for the purpose of music making. The core of HBB is instru-

mental mimicry: typically, the first sounds a beatboxer learns are those that reproduce the drum set sounds, i.e. kick, hi-hat, snare/rimshot, cymbal. More experienced beatboxers, however, can be considered as multivocalists, as they exploit a wide variety of vocal techniques such as rapping, singing, overtone singing, scratching, etc. depending on the style of music they want to produce. In particular, the humming technique can be used to give the impression of multiple sound sources within the same beatboxer: a rhythmic line and a melodic line can be produced simultaneously. This technique is well known by beatboxers and is generally resumed as the technique that allows a beatboxer to produce multiple sounds at the same time using the air present in the mouth to produce the rhythm, and the voice to produce the melody. However, how this is achieved remains mostly unexplored from a scientific standpoint. This study focuses on three categories of drum sounds (kick, hi-hat, rimshot) produced as regular HBB sounds or in the humming technique. Some studies have shown that regular kick, hi-hat, and rimshot are generally produced via a piston-like action of the closed glottis [3, 6, 1], i.e. using a glottalic initiation mechanism [4]. The only published study so far that directly investigates humming boxemes (i.e. HBB sounds) has shown that the humming versions of these three boxemes are produced via a pushing or pulling action of the tongue [5], i.e. using a velaric initiation mechanism [4]. The present study aims at elucidating the similarities and differences in terms of breathing strategy and articulatory mechanism between regular and humming kick, hi-hat, and rimshot as well as giving some insights on how the vocal tract is configured when producing different sounds simultaneously.

## II. METHODS

Four male French speaking beatboxers (S02-S05) were recorded, three professionals and one amateur, aged 20 to 38 years. The recordings took place in the semi-anechoic room of GIPSA-lab in Grenoble, a place of biomedical research authorized by the ARS Auvergne-Rhône-Alpes. Electromagnetic articulography (EMA WAVE, NDI, Canada) and respiratory inductance plethysmography (RIP, ETISENSE, France) were combined with electroglottographic, acoustic and video recordings. Audio and EGG signals were sampled at 20 kHz, RIP and EMA signals at 200 Hz.

The three boxemes kick (**P**), hi-hat (**t**), rimshot (**K**) were produced 12 times each in a row. Each repetition was preceded by [sæsɛlə] (English translation: “this is the”). Each boxeme sequence was produced in regular HBB, then in the humming technique. Further, the *beat*, i.e. a musical phrase, **PtKtPtKt** was repeated 10 times as regular HBB, and 10 times as humming HBB. The data recorded by the different systems were all synchronized together. The audio recordings were manually segmented and annotated using Praat software [2]. Spatial trajectories of 9 coils placed on five flesh points of the tongue (apex/blade, middle, right, left and dorsum) and four flesh points of the lips (upper and lower, median and lateral) were extracted from the EMA recordings. Mean trajectories and variance were computed using the commercial software package MATLAB.

## III. RESULTS

The three professional beatboxers (S03, S04, S05) produced two versions of the humming boxemes: one was only a sequence of drum sounds, i.e. the rhythmic line (**RL**), the other was a superposition of drum sounds (**RL**) and a hummed melodic line (**ML**). The presence of vocal-fold vibration is attested by the EGG signal in Fig. 2. However, one beatboxer alternated vocal-fold vibration and glottal stops when producing his **ML**. The amateur beatboxer (S02) gave only one version of humming boxemes as post-voiced boxemes: no vocal-fold vibration occurred synchronously to the drum-sound production, but was present right after. No vocal-fold vibration was detected during regular HBB production.

Breathing strategies varied among beatboxers and stimuli. Shorter tasks such as boxeme repetition held the most variability. However, a typical pattern emerged during humming **RL** tasks: an increase in thoracic and abdominal circumferences before the initiation of the *beat* was followed by an alternation of decrease and increase during the *beat*. Fig. 2 shows that this alternation can be similar to breathing behavior at rest, but was not related to the acoustic outcome. When voicing was added during humming **RL+ML** executions, the evolution of thoracic and abdominal circumferences was sim-

ilar to speech: an increase before vocal-fold vibration initiation attested of air intake, followed by a regular decrease during voicing.

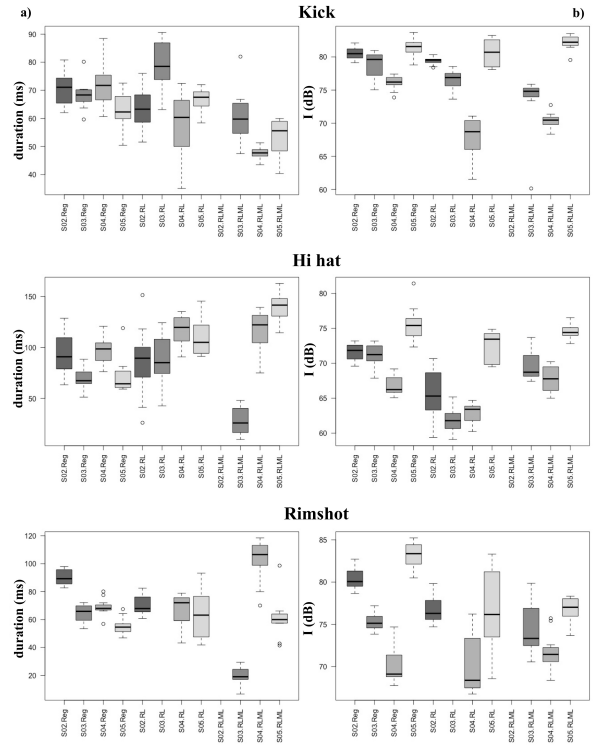


Figure 1: Distribution of a) sound duration (in ms) and b) sound intensity (in dB) for each boxeme produced by each beatboxer (S02-S05) as regular HBB (Reg), humming (RL), and voiced humming (RLML).

Regular HBB production showed the most varied breathing strategies among beatboxers, especially for shorter tasks (boxemes repetition). Longer tasks, more similar to real-life HBB, attested of a typical behavior: a tendency towards stabilization of thoracic (and possibly abdominal) circumference during the *beat* execution, with small local variations in correspondence with each boxeme acoustic production. In the case illustrated in Fig. 2, the more prominent local variations occurred in correspondence with **K** and indicated a rapid increase in thoracic circumference suggesting a small inhalation during the boxeme production.

Articulatory behavior was quite consistent among the four beatboxers for the realization of **P** and **t**, whereas **K** showed more variability. Fig. 3 illustrates mean trajectories and acoustic outcomes for S04. **P** was always produced as a bilabial occlusive. However, the tongue was particularly active in the realization of the three variants (Regular, humming **RL**, and humming **RL+ML**). As for the two humming versions, the articulatory data show

that the superposition of the **ML**, in this case the vocal-fold vibration, to the **RL** did not impact the lingual or labial movements. The tongue was raised high against the palate in the back of the oral cavity and was pushed forward right before and during the occlusion release and the acoustic realization. Breathing data showed no relation between breathing and acoustic production of **P**. In the case of the regular **P**, the tongue assumed a lower position in the oral cavity and underwent an upward displacement that began before the occlusion release and ended after the cessation of the sound. Breathing data showed local decrease in thoracic circumference around sound production, suggesting the use of an egressive airstream. This resulted in slightly shorter and softer humming sounds compared to their regular equivalents (Fig. 1). **t** was always produced as an alveolar occlusive. Two main articulatory strategies were observed for the humming versions. One strategy consisted in holding the tongue against the palate and then suddenly producing a rapid downward movement, especially in the middle region, during which the occlusion was released and the sound took place. The other (shown in Fig. 3) was via occlusion of the vocal tract in the anterior and posterior region of the oral cavity, creating an air pocket between the middle region of the tongue and the palate and subsequently compressing the trapped air via a pushing action of the middle section of the tongue, releasing the anterior occlusion and producing the sound. In both articulations the tongue assumed a high position, especially in the back region of the oral cavity. The breathing data showed no relation with sound production. Regular **t** was achieved with a lower position of the tongue. Only the more anterior part of the tongue made contact with the palate during the occlusion phase. At occlusion release, the anterior portion of the tongue was lowered and at the same time the posterior portion of the tongue underwent an upward movement. Breathing data showed local decrease in thoracic circumference, suggesting the use of an egressive airstream. The humming versions of **t** generally were longer and softer than the regular **t**. **K** was achieved using the most different articulatory strategies among the beatboxers. For the most part, the humming versions were realized pushing the tongue against the palate and then pulling it down in a rapid motion, while the more anterior region of the tongue was kept in contact with the palate (Fig. 3) and the occlusion was released on one side of the tongue. In the humming **PtKt** task, one beatboxer also produced **K** as a bilabial occlusive, where the pressure buildup was achieved via compression of the cheeks. Again, no relation emerged between breath and acoustic realization. Regular **K** was realized in two different ways. An occlusion was created in the back region of the tongue against the palate, then released via a rapid downward motion of

the tongue (Fig. 3). In the regular **PtKt** task, two beatboxers used a different articulatory behavior with a different acoustic outcome: the tongue was kept in contact with the palate during the occlusion phase, then the occlusion was released in the back region of the tongue, while the front portion of the tongue was kept in contact with the palate. Breathing data showed a local increase in thoracic circumference, suggesting the use of an ingressive airstream. Once again, humming boxemes resulted as softer and generally slightly shorter than their regular equivalent.

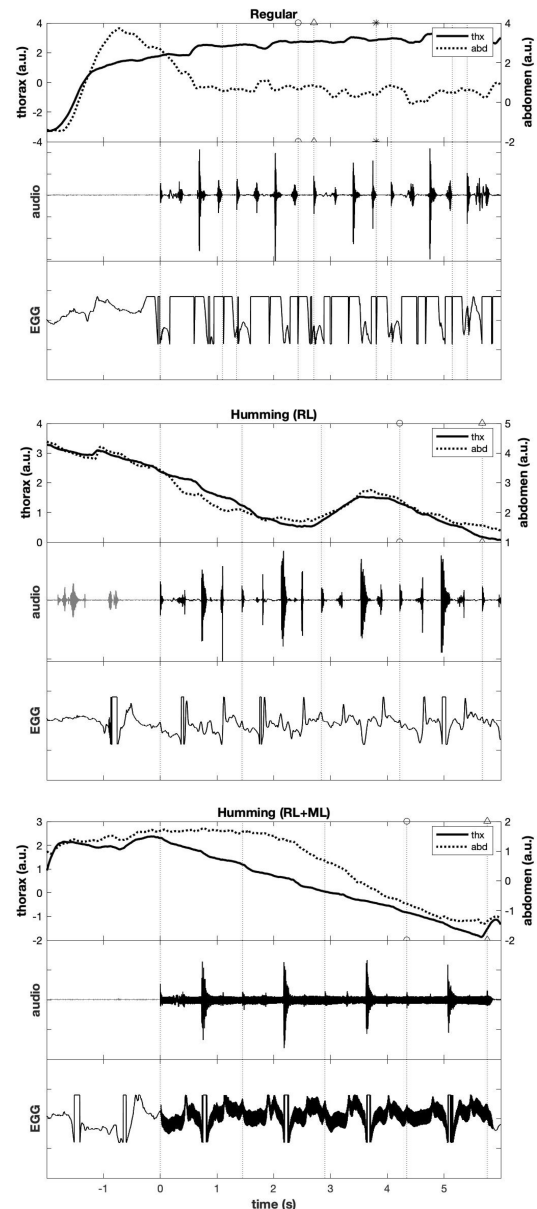


Figure 2: Breathing, audio, and EGG signals of S04 producing the *beat PtKt*. y-axes are arbitrary scales.

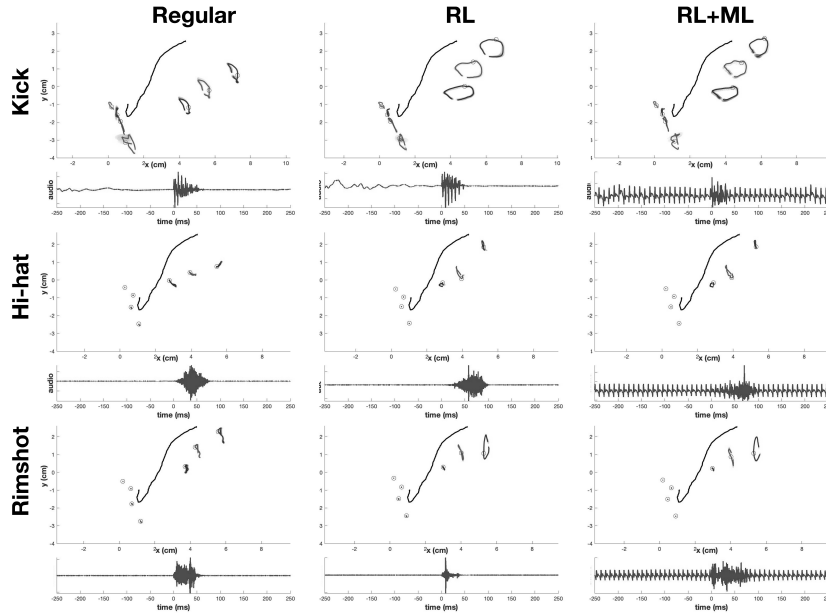


Figure 3: Articulatory trajectories of tongue and lip coils in the mid-sagittal plane during regular beatboxing and humming without (RL) or with (RL+ML) melodic line, for singer S04. For each sequence, audio signal of a representative token is plotted. Solid line: palate contour

#### IV. DISCUSSION

Beatboxers naturally produced two humming versions of **P**, **t**, and **K**: one was the **RL** without **ML**, the other both **RL** and **ML**. We found that the term “humming HBB” does not imply the presence of a **ML**, but rather the choice of a particular articulatory strategy for the **RL** that is restrained to the oral cavity. This study showed that, while for regular **P**, **t**, and **K** breathing and articulatory behavior are related, with a likely glottalic or pulmonic initiation mechanism in most cases, the humming equivalents systematically switch to a velaric (mostly lingual) egressive or ingressive initiation mechanism. This bares two main consequences: on the one hand, humming boxemes are generally less intense than regular HBB boxemes; on the other hand, the use of an oral airstream to produce the **RL** allows for the dissociation of breathing and articulation. The high position of the back of the tongue divides the vocal tract into two functional sections that can produce two different sounds at the same time.

#### V. CONCLUSION

In humming HBB, the synchronous production of a rhythmic line and a melodic line is achieved by isolating the oral cavity from the rest of the vocal tract. The oral cavity functions on its own to produce the rhythmic line. Humming kick **P**, hi-hat **t**, and rimshot **K** are produced via velaric initiation mechanisms.

This leaves the upstream part of the vocal tract (laryngeal and pharyngeal spaces) available for breathing or producing the melodic line. In the latter case, the hum-

ming sound source generated by vocal-fold vibration is propagated into the nasal cavities. This is a skilful and original use of the vocal tract, regularly performed by beatboxers.

#### V. REFERENCES

- [1] R. Blaylock, N. Patil, T. Greer, and S. S. Narayanan. Sounds of the human vocal tract. In *Proceedings of Interspeech*, pages 2287–2291, 2017.
- [2] P. Boersma. Praat: doing phonetics by computer. <http://www.praat.org/>, 2006.
- [3] T. De Torcy, A. Clouet, C. Pillot-Loiseau, J. Vaissiere, D. Brasnu, and L. Crevier-Buchman. A video-fiberscopic study of laryngopharyngeal behaviour in the human beatbox. *Logopedics Phoniatrics Vocology*, 39(1):38–48, 2014.
- [4] P. Helgason. Sound initiation and source types in human imitations of sounds. In *Proceedings of FONETIK*, pages 83–88, 2014.
- [5] A. Paroni, N. Henrich Bernardoni, C. Savariaux, H. Lœvenbruck, P. Calabrese, T. Pellegrini, S. Mouysset, and S. Gerber. Vocal drum sounds in human beatboxing: An acoustic and articulatory exploration using electromagnetic articulography. *The Journal of the Acoustical Society of America*, 149(1):191–206, 2021.
- [6] M. Proctor, E. Bresch, D. Byrd, K. Nayak, and S. Narayanan. Paralinguistic mechanisms of production in human “beatboxing”: A real-time magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, 133(2):1043–1054, 2013.