Human Beatboxing : A Multi-Instrumental Pilot

Alexis Dehais Underdown[1], Paul Vignes[1], Lise Crevier Buchman[1,2], Didier Demolin[1]
[1]Laboratoire de Phonétique et Phonologie, UMR 7018 CNRS-Sorbonne Nouvelle, [2]Hôpital Foch, Univ. VSQ

ABSTRACT

This comparative study aims to describe laryngeal, acoustic and aerodynamic characteristics of instrumental imitations that is the beatboxed classic kick drum [p'] and snare drum [p͡f'] (PF-snare) by one artist. Beatboxed sounds were produced in isolation and in beatboxed patterns. Aerodynamic (intraoral pressure, oral airflow, nasal airflow), acoustic, electroglottographic and laryngoscopic data were acquired. Based on the acoustic and aerodynamic data we discuss the coordination of the articulators and the planification of articulatory commands of the classic kick drum and the PF-snare drum.

GOAL

1) Description of [p'] and [pf']:
  → Aerodynamic characteristics
  → Laryngeal articulation
  → Spectral components
2) Different contexts of production
  → Produced in isolation
  → Production in a beat pattern
  (i.e. with other beatboxed sounds)

References

- Blaylock, R., Patil, N., Greer, T., Narayanan, S. (August, 2017). Sounds of the Human Vocal Tract. *INTERSPEECH 2017*, (pp. 2287-2291). Stockholm, Sweden.
- Dehais Underdown, A., Crevier Buchman, L. & Demolin, D. (2019). Acoustico-Physiological coordination in the Human Beatbox: A pilot study on the beatboxed Classic Kick Drum. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 142-146). Canberra, Australia: Australasian Speech Science and Technology Association Inc.: Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.).
- Edmondson, J. A., & Esling, J. H. (2006). The valves of the throat and their functioning in tone, vocal register and stress: laryngoscopic case studies. *Phonology, 23(2)*, 157-191
- Ghio, A., & Teston, B. (2004). Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *International Conference on Voice Physiology and Biomechanics*, (pp. 55-58). Marseille, France.
- Patil, N., Greer, T., Blaylock, R., & Narayanan, S. . (2017). Comparison of Basic Beatboxing Articulations between Expert and Novice Artists using Real-Time Magnetic Resonance Imaging. *INTERSPEECH 2017*, (pp. 2277-2281). Stockholm, Sweden
- Proctor, M., Bresch, E., Byrd, D., Nayak, K. & Narayanan, S. (2013). 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.

Methods

Beatboxer: 1 beatboxer, 35 y. o (see Dehais Underdown, Crevier Buchman & Demolin, 2019 for further details)
Corpus: Production of his beatboxing repertoire, sounds in isolation and in Beat Patterns
Instrumentation:
→ Aerodynamics : intraoral pressure (i.e. Po) only for labials + Oral airflow (i.e. Oaf) + Nasal Aifrlow (i.e. Naf) + Electroglottography (i.e. EGG) => EVA 2 workstation (SQLab-LPL, Aix-en-Provence, France, cf. Ghio & Teston, 2004)
→ Audio: Acoustic Waveform synchronized with aerodynamic signals
→ Laryngeal Nasofibroscopy: Kay-Pentax®FNL10RP) with a DigitalStrobe®, RLS91000 (Kay Elemetrics, Lincoln Park, NJ, USA) => 25 fps
Analysis:
→ Aerodynamics : total duration (ms) + closure duration + friction duration (for affricates) + peak of Po (hPa) + peak of Oaf (dm^3/s) + peak of Naf (cm^3/s) + volume of air (cm^3) (i.e. integral of airflow signals) + impressionistic description of EGG signal.
→ Audio: FFT spectrum (25ms window) + Center of Gravity (CoG) + Skewness + Kurtosis
→ Laryngeal Nasofibroscopy: Analysis of contraction & dilatation of the laryngeal valves (Edmonson & Esling, 2006)

Results

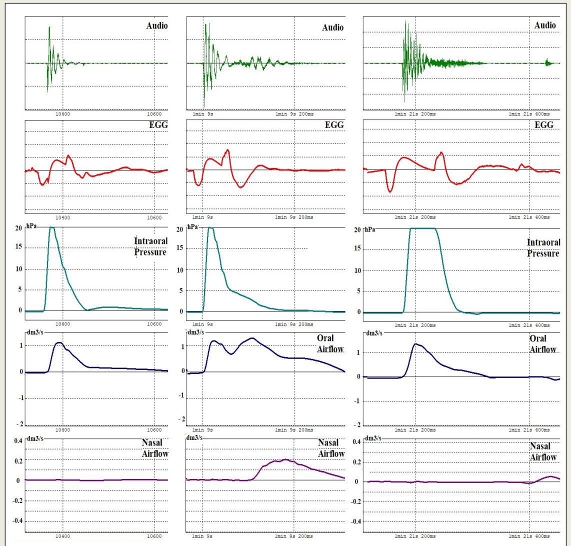

Figure 1: Acoustic waveform, EGG signal, Po (hPa) and Oaf and Naf (dm^3/s) of [p'] (left) [p'ɸ] (middle) and [pf'] (right).

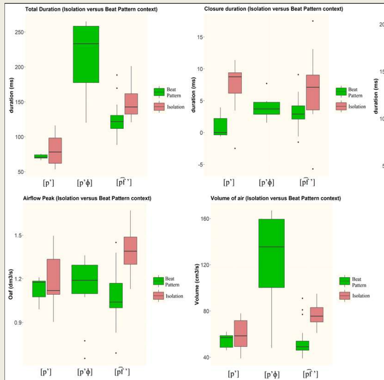

Figure 2: Box plot of total duration (top left), closure duration (top middle), frication duration (top right), airflow peak (bottom left) and volume of air (bottom right) for [p'], [p'ɸ] and [p͡f'] produced in isolation (red) and in a beat pattern (green)

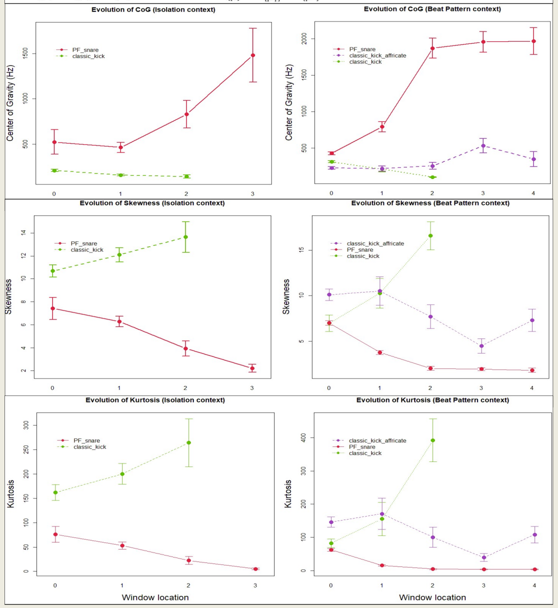

Figure 3: Compared evolution of CoG (top), Skewness (middle) and kurtosis (bottom) of the PF-snare (red), classic kick (green), affricated kick (purple) in isolation (left) and in a beat pattern (right).

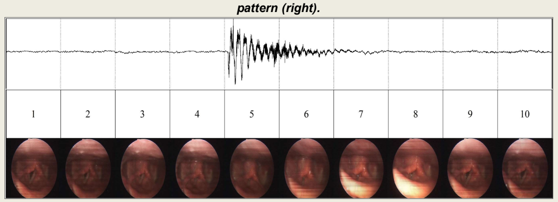

Figure 4: Acoustic waveform and laryngoscopic images of [pf']. Frame 1 to 3 = glottal adduction, frame 4 to 6 = supraglottal compression of arytenoid cartilages and the aryepiglottic folds, frame 7 to 9 = glottal abduction involving vocal folds and ventricular folds

Discussion

→ Different patterns of articulatory coordination => [p'] & [pf'] are not sequentially produced but synchrounously (i.e. (quasi)simultaneous closure + release)
  → Stops (including ejectives) are produced sequentially across languages

→ /p'/ occurs in languages but */pf'/ does not
  → /p'/ & /pf/ are rare so [pf'] is not likely to be phonemic in languages

→ short closure time (> 10 ms) => lips closure > velopharyngeal closure ?
  → MRI data ? (cf. Proctor et al. 2013; Blaylock et al. 2017; Patil et al. 2017)

→ 2 different lip postures at the end of [p']:
(1) [p'] lips = "projected" forward, no 2nd constriction;
  → Low CoG + high skewness (low frequency components) and high kurtosis (flat spectrum)
(2) [pf'] = the lower lip + jaw move backward to meet the upper teeth => 2nd constriction.
  → Increasing CoG + low skewness (high frequency components) and low kurtosis (peaked spectrum)

→ Temporal reduction for both [p'] and [pf'] when they are produced in pattern with other sounds => further work should focus on metric and rhythm and its impact on articulation

→ [p'] has two realization in the data :
(1) [p'ɸ] = bilabial voiceless glottalic egressive affricate
(2) [p'ⁿ] = bilabial voiceless glottalic egressive nasalized stop
  → They seem to be the result of the interaction between respiration and articulation => « disconnection » of oral tract and nasal tract
  → Further work on breathing pattern in human

Conclusion

→ HBB = musical language ≠ spoken languages
  → No semantic components in HBB
  → Removing linguistic constraints may change the use of the vocal tract and articulators recover all their degree of freedom
→ HBB may allow us to know more about vocal tract articulatory capacities
  → Implication for general phonetics (i.e. diversity of phonological systems) and clinical phonetics (i.e. fun exercises for children)

Acknowledgements

We thank the participant for his time, the ASA for funding the travel.