SPEED RATE EFFECTS ON HUMAN BEATBOXING: AN AERODYNAMIC STUDY

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]Unité Voix, Parole Déglutition, Service ORL et de Chirurgie de la Face et du Cou, Hôpital Foch
alexis.dehais-underdown@sorbonne-nouvelle.fr

ABSTRACT

The present study aims to describe speed rate effects (i.e. tempo) on Human Beatboxing based on aerodynamic data. Intraoral pressure (Po), oral airflow (Oaf), nasal airflow (Naf) and the acoustic signal were simultaneously recorded. The experiment consisted in the production of various Beat Patterns composed of *kick drums*, *hi-hats* and *snare drums*. They were produced at different rates: 90 beats per minute (BPM), 120 BPM and 150 BPM. Based on Po and Oaf, we estimated the relative constriction area. Results show temporal reduction of patterns ($\rho = -0.97$) and sounds ($\rho = -0.37$). Subjects tended to decrease intraoral pressure and increase constriction area as speed rate increased. In few cases, subjects increased intraoral pressure and decreased constriction area as rate increased. Human Beatboxing constitute an original contribution to discuss production mechanisms with non-linguistic data.

Keywords: Experimental Phonetics, Beatboxing, Reduction, Tempo, Aerodynamics

1. INTRODUCTION

The present study aims to describe speed rate (i.e. tempo) effects on Human Beatboxing (HBB) based on aerodynamic data. Speech rate is known to be a factor of reduction affecting supralaryngeal and laryngeal gestures. During fast speech, articulators' velocity is increased and consequently gestures' duration and amplitude are reduced [1, 2, 3, 4, 5]. At fast rates, the interval of time between gestures is shortened, the degree of overlap may increase as well [4, 5]. Speech rate was found to have little or no effects on intraoral pressure and oral airflow peaks [6, 7, 8, 9]. Though, these studies do not account for individual patterns and only investigate pulmonic egressive obstruents. HBB is a musical technique produced with vocal tract movements. A study of [10] found that tempo affects the duration of beatboxed sounds and patterns.

Aerodynamics is a branch of fluid mechanics dealing with the motion of gases. During vocal production (e.g. speech, singing), vocal tract volume continuously changes over time and generates different pressure levels and different types and levels of airflow (i.e. volume velocity). From intraoral pressure and airflow measurements, we can estimate particle velocity and the relative constriction area [11]. *Particle velocity* (v) is the rate of change in particle position. Particle velocity depends on several factors as pressure, air density or gravitational acceleration [12]. Catford [11] offers a simplified equation for phonetic purposes: $v = 412\sqrt{P}$, where P is the measured intraoral pressure. *Relative constriction area* (A) is derived from volume velocity and particle velocity $A = U/v$ where U is the measured oral airflow and v the estimated *particle velocity*.

To investigate speed rate effects on HBB, subjects were asked to produce *Beatboxing Patterns* at 3 different tempo: 90, 120 and 150 beat per minutes (BPM). Tempo is a musical term that refers to the interval of time separating pulsations (e.g. interval of time separating the metronome's beats) and defines how "quick" a musician plays. Low tempo are characterized by long intervals between notes and fast tempo are characterized by shorter intervals. In terms of production, at fast tempo, musician will have less time to reach the targets (e.g. key strike for a pianist). Concerning HBB, we hypothesize that subjects will have less time to reach the targets as tempo increases and we expect intraoral pressure to decrease and constriction area to increase as tempo increases.

2. METHODS

2.1. Data Acquisition

Aerodynamic and acoustic signals were acquired with EVA2 Workstation [13]. EVA2 Workstation allows simultaneous recording of acoustic signal, intraoral pressure (hPa, $1hPa = 1.02cmH_2O$), oral airflow and nasal airflow (dm^3/s). The acoustic waveform was obtained by EVA2 integrated microphone. Oral Airflow (Oaf) was collected using a flexible silicone mask pressed on the subjects' mouth. Nasal Airflow (Naf) was obtained through a tube placed in the nostril. Intraoral pressure (Po) was obtained inserting a small tube into the pharynx through the nasal cavity. Whenever Po and Oaf clipped, tokens were excluded from the analysis.

2.2. Protocol and Corpus

The data presented here was excerpted from a larger database of aerodynamic data from 4 professional beatboxers: 3 males (VP, CJ, GA) and 1 female (AI). Most of them have 12 or 13 years of training except AI that has only 4 years of training. All of them won at least one french championship and two of them won a world championship. The experiment consisted in the production of various Beat Patterns (BP) produced at 90 BPM, 120 BPM and 150 BPM repeated 4 times each. Tempo was given through a vibrating metronome placed and the subjects' wrist.

The corpus is composed of 11 patterns with the same metrical, rhythmical and melodic structure but with different phonetic structure. Figure 1 gives an example structure of the basic pattern [p' t͡s' ↓k͡ʟ t͡s' p' p' t͡s' ↓k͡ʟ t͡s']. The same pattern may also be produced with a different phonetic structure such as [t̪' t͡s' t͡ʃ' t͡s' t̪' t̪' t͡s' t͡ʃ' t͡s'] for example. On the Figure, the first line shows a 90BPM metronome's pulsation, the second line shows the melodic structure with drums alternations, the third line indicates the position in the metrics and the fourth line gives a transcription of the phonetic structure. The analysis focuses on the pattern illustrated in Figure 1. In this BP, [p'] was produced as an ejective bilabial stop. [t͡s'] was produced as an ejective dental affricate by subjects VP, AI, CJ and as a pulmonic egressive dental affricate [t͡sː] in position 9 by subject GA. [↓k͡ʟ] was produced as a pulmonic ingressive velar lateral affricate by subjects VP, CJ, GA and as an implosive velar stop and a pulmonic ingressive velar lateral fricative [↓k̊͡ʟ] by subject AI. In total, 48 patterns and 396 tokens were analyzed.

2.3. Aerodynamic measures

For each sound, intraoral pressure (Po) and oral airflow (Oaf) peaks were automatically extracted in Praat. Then, we calculated particle velocity $v = 412\sqrt{P}$, where P is the measured pressure, and relative constriction area $A = U/v$, where U is the measured oral airflow. We also extracted

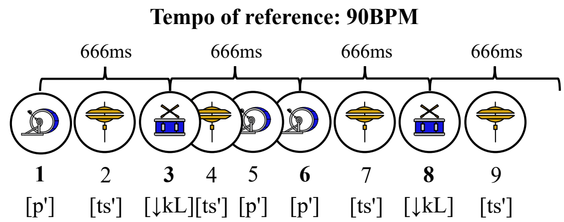

Figure 1: Metrical structure of a Beat Pattern at 90BPM. Line 1: metronome's pulsation, line 2: melodic structure, line 3: metrical position, line 4: phonetic structure

sound duration and patterns duration to assess temporal reduction of beatbox patterns and sounds. Sound duration was measured from the beginning of the occlusion (when Po starts increasing) to the offset of the acoustic signal. Pattern duration was measured from the beginning of the first sound to the end of the last one. Based on patterns duration, we calculated the actual tempo at which subjects produced patterns. Spearman's coefficient (ρ) were calculated in R between temporal and aerodynamic variables. Sounds were analyzed for each participant in each position to investigate individual behaviors and possible effect of sounds and metrical positions regarding speed rate effects.

3. RESULTS

Spearman's coefficients showed very high correlation between the tempo of reference (i.e. 90, 120, 150BPM) and BP duration ($\rho = -0.97$). BP duration decreases when speed rate increases. Correlation between the tempo of reference and sound duration is lower ($\rho = -0.37$). Sounds are less affected than BPs, suggesting inter-gestural intervals decrease more than sounds themselves. There was no correlation between tempo and aerodynamic measures.

3.1. Speed rate effects on pressure

Intraoral pressure changes as a function of speed rate allows us to infer changes of the initiatory gesture (i.e. gesture setting air in motion). A decrease of the pressure would suggest an undershoot of the gesture (i.e. less volumetric reduction) while an increase of the pressure would suggest an overshoot of the gesture (i.e. more volumetric reduction). Figure 2 shows a scatter-plot of intraoral pressure as a function of tempo. There is a tendency of intraoral pressure to decrease at fast rate. Though, some exceptions and non-linear effects are observed.

There is an increase in intraoral pressure as speed rate increases for subject CJ in position 6 ($\rho = 0.52$)

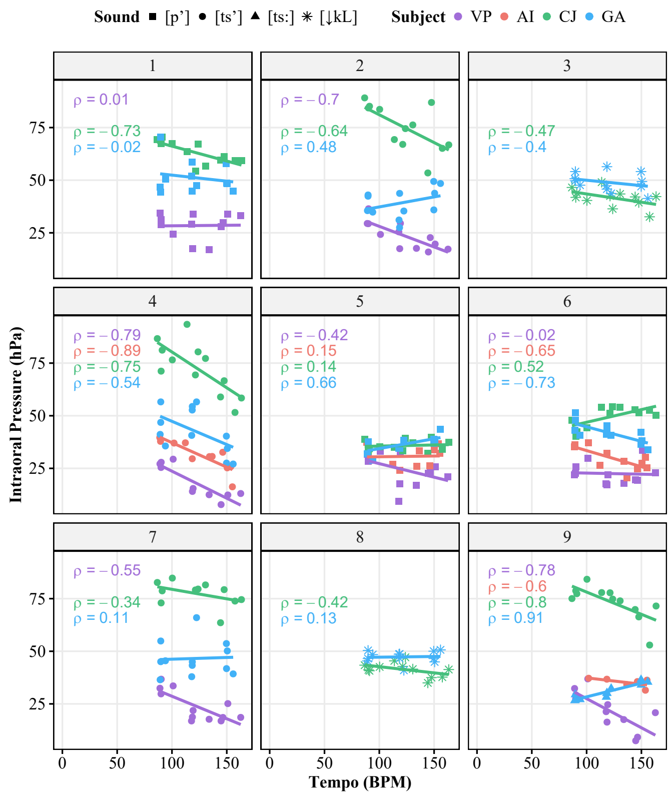

Figure 2: Intraoral pressure as a function of tempo. Each panel represents one position.

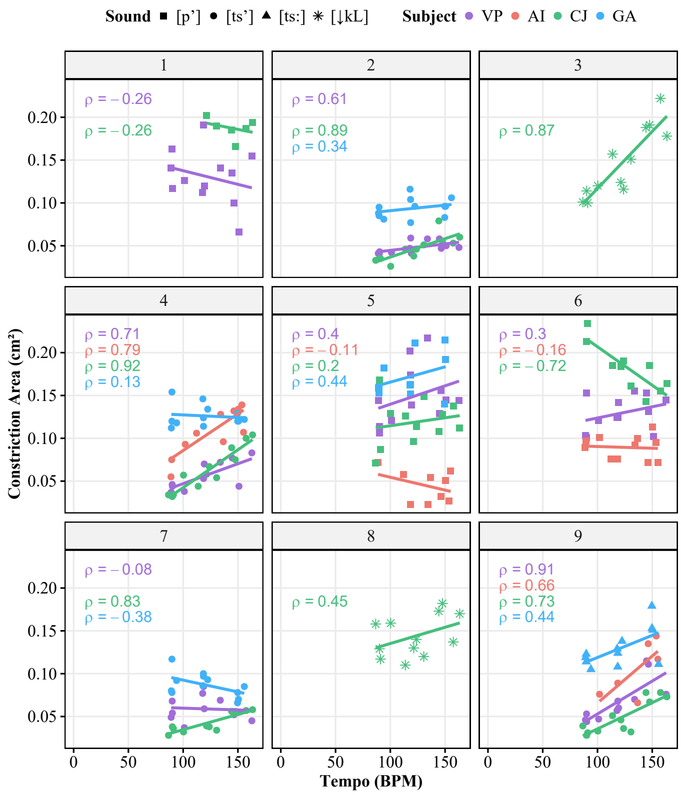

Figure 3: Constriction area as a function of tempo. Each panel represents one position.

and for subject GA in position 5 ($\rho = 0.66$) and 9 ($\rho = 0.73$). This increase suggest that there is a greater volume reduction as tempo increases. In some cases, initiation remains stable as for example subject VP in position 1 and 6.

There are non-linear effects of speed rate on initiation. For example, if we look at the data from CJ in position 6, Po increases between low tempo (i.e. 90BPM) and medium tempo (i.e. 120BPM) but remains stable between medium and fast (i.e. 150BPM). Another non-linear effect can also be observed for subject GA in position 4 where Po increases at 120BPM but decreases at 150BPM.

There is a consistent tendency across subjects in position 4, suggesting position in the metric influences reduction. Indeed, intraoral pressure reduces for all subjects concerning [t͡s'] in position 4 ($\rho < -0.54$). In position 9, there is an important decrease of the pressure for all subjects except for GA who produced [t͡sː] with increasing pressure ($\pm 10hPa$). Less effect of rate on pressure is observed for [p'] and little effect on pressure is observed for [↓k͡ʟ]. Concerning [p'] in position 1, CJ decreases the pressure about 16hPa at fast rate while for VP and GA, pressure is more or less stable. In position 5, there is a decrease of the pressure for VP. In position 6 we observe a decrease of presure for AI and GA and an increase for CJ.

3.2. Speed rate effects on constriction area

The estimation of relative constriction area allows us to infer changes in the articulation. The tendency observed in the data is an increase of the constriction area as speed rate increases. In few instances, we observe a decrease of the constriction area as for subject VP in position 1, subject CJ in position 6 (i.e. [p']) and GA in position 7 (i.e. [t͡s']). In few cases, constriction area remains stable (e.g. VP position 7, AI position 6, GA position 4). Non-linear effects are also observed concerning speed rate effects on constriction area. For example, [t͡s'] in position 2, GA shows an increase of the constriction area at 120BPM and a decrease from 120 to 150BPM.

Once again, we observe a consistent tendency across subjects in position 4 and 9. Indeed, the data suggest a larger constriction area for all subjects, except GA in position 4 where constriction remains stable as speed rate increases. In initial position, VP and CJ show a decrease of the constriction area. A similar observation is made from CJ's data in position 6. While pressure remained stable for the *inward k-snare* [↓k͡ʟ], data from CJ shows an increase of the constriction area as tempo increases.

4. DISCUSSION

The data showed temporal reduction of beatboxing patterns ($\rho = -0.97$) and sounds ($\rho = -0.37$). Sounds decreased their duration to a lesser extent suggesting a greater effect on intergestural duration. As hypothesized, when rate is increasing, subjects tended to decrease intraoral pressure and increase constriction area. It was found that *hi-hats* produced in position 4 and 9 were the most affected by increased tempo. Contrary to *kick drums* and *snare drums*, *hi-hats* are not produced in phase with the metronome beat therefore, they constitute weaker positions. Furthermore, position 4 forms a complex interval along with the preceding snare and the following kick drum as shown on Figure 1. Thus, it is not surprising to observe reduction in an interval where multiple targets are to be reached (i.e. velar, dental and bilabial).

Our data also shows cases where the pressure is increased and the constriction area is decreased. These data contradicts our initial hypothesis. We think that timing (i.e. synchronization of the production with the metronome), may help explaining this contradiction. Indeed, Figure 2 and 3 shows that subjects are beatboxing more quickly and slowly than the tempo of reference. This suggests that the timing of production differ from the metronome's pulsation. Measuring and quantifying the deviation from the reference of the metronome for each interval would discard or confirm possible effects of timing on HBB production. An acoustic study is needed to investigate if changes are observed in the acoustic output.

5. CONCLUSION

The data presented here shows an effect of tempo on production of HBB. Though, future studies are needed to understand individual adaptation to increasing rate, particularly to explain the contradictions in our data. A study on timing could be relevant to further understand if timing interacts with tempo and if subjects try to compensate tempo and timing effects (i.e. overshoot at fast tempo). Such study would constitute an original contribution to discuss production mechanisms and motor control with non-linguistic data.

6. ACKNOWLEDGMENT

This physiological protocol was approved by an ethical committee (RCB-ID n° 2020-A00246-33). The research is supported by the Delegation of Research and Innovation of the Hôpital Foch and

LABEX EFL (Empirical Fundation of Linguistics, ANR-10-LABX-0083).

7. REFERENCES

[1] B. Lindblom, “Spectrographic study of vowel reduction,” *The journal of the Acoustical society of America*, vol. 35, no. 11, pp. 1773–1781, 1963.
[2] T. Gay, T. Ushijima, H. Hiroset, and F. S. Cooper, “Effect of speaking rate on labial consonant-vowel articulation,” *Journal of Phonetics*, vol. 2, no. 1, pp. 47–63, 1974.
[3] D. J. Ostry and K. G. Munhall, “Control of rate and duration of speech movements,” *The Journal of the Acoustical Society of America*, vol. 77, no. 2, pp. 640–648, 1985.
[4] K. Munhall and A. Löfqvist, “Gestural aggregation in speech: Laryngeal gestures,” *Journal of Phonetics*, vol. 20, no. 1, pp. 111–126, 1992.
[5] D. Byrd and C. C. Tan, “Saying consonant clusters quickly,” *Journal of Phonetics*, vol. 24, no. 2, pp. 263–282, 1996.
[6] H. J. Arkebauer, T. J. Hixon, and J. C. Hardy, “Peak intraoral air pressures during speech,” *Journal of Speech and hearing Research*, vol. 10, no. 2, pp. 196–208, 1967.
[7] T. Hixon, “Turbulent noise sources for speech,” *Folia Phoniatrica et Logopaedica*, vol. 18, no. 3, pp. 168–182, 1966.
[8] W. Brown Jr and R. E. McGlone, “Relation of intraoral air pressure to oral cavity size,” *The Journal of the Acoustical Society of America*, vol. 45, no. 1, pp. 323–323, 1969.
[9] A. Malécot, “The effect of syllabic rate and loudness on the force of articulation of american stops and fricatives,” *Phonetica*, vol. 19, no. 4, pp. 205–216, 1969.
[10] A. Dehais-Underdown, P. Vignes, L. C. Buchman, and D. Demolin, “Human beatboxing: A preliminar study on temporal reduction,” in *12th International Seminar on Speech Production*, 2020.
[11] J. C. Catford, *Fundamental problems in phonetics*. Bloomington And London: Indiana University Press, 1977.
[12] C. H. Shadle, “The aerodynamics of speech,” *The Handbook of Phonetic Sciences*, pp. 39–80, 2010.
[13] A. Ghio and B. Teston, “Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies,” in *International Conference on Voice Physiology and Biomechanics*, 2004, pp. 55–58.